Journal of Biomolecular NMR, 7 (1996) 331–334
ESCOM

J-Bio NMR 362

Local helix content in an alanine-rich peptide as determined by the complete set of $^3J_{HN\alpha}$ coupling constants

Glenn L. Millhauser[a,*], Chris J. Stenland[a], Kimberly A. Bolin[a] and Frank J.M. van de Ven[b,**]

[a]*Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, U.S.A.*
[b]*Biophysical Chemistry Laboratory, Nijmegen SON Research Center, University of Nijmegen, NL-6525 ED Nijmegen, The Netherlands*

Received 24 April 1996
Accepted 29 April 1996

Keywords: J coupling; Coupling constants; Inverse fitting; Alanine; Helix; Peptide

Summary

Alanine-rich peptides serve as models for exploring the factors that control helix structure in peptides and proteins. Scalar $C^{\alpha}H$-NH couplings ($^3J_{HN\alpha}$) are an extremely useful measure of local helix content; however, the large alanine content in these peptides leads to significant signal overlap in the $C^{\alpha}H$ region of ^1H 2D NMR spectra. Quantitative determination of all possible $^3J_{HN\alpha}$ values is, therefore, very challenging. Szyperski and co-workers [(1992) *J. Magn. Reson.*, **99**, 552–560] have recently developed a method for determining $^3J_{HN\alpha}$ from NOESY spectra. Because $^3J_{HN\alpha}$ may be determined from 2D peaks outside of the $C^{\alpha}H$ region, there is a much greater likelihood of identifying resolved resonances and measuring the associated coupling constants. It is demonstrated here that $^3J_{HN\alpha}$ can be obtained for every residue in the helical peptide Ac-$(AAAAK)_3A$-NH_2. The resulting $^3J_{HN\alpha}$ profile clearly identifies a helical structure in the middle of the peptide and further suggests that the respective helix termini unfold via distinct pathways.

Designed helical peptides, often rich in alanine, serve as models for exploring helix-stabilizing influences in proteins. Over the last several years there has been ongoing interest in determining the distribution of the helix content throughout individual peptide sequences (Osterhout et al., 1989; Bradley et al., 1990; Lyu et al., 1990; Merutka and Stellwagen, 1990; Chakrabartty et al., 1991; Liff et al., 1991; Storrs et al., 1992; Fiori et al., 1993; Merutka et al., 1993; Miick et al., 1993; Rohl and Baldwin, 1994; Zhou et al., 1994). NMR is well-suited for such studies. Hydrogen exchange, sequential NOEs and $C^{\alpha}H$-NH scalar couplings ($^3J_{HN\alpha}$) all report on the helical nature of individual residues (Wüthrich, 1986). However, in peptides with a high alanine content, NMR signal overlap often obscures the desired information. To address this challenge, researchers have used several approaches, including isotopic labeling (Rohl and Baldwin, 1994), enhancement of the helix content with covalent side-chain linkages (which can result in a greater signal dispersion) (Zhou et al., 1994) and the selective incorporation of non-Ala residues (Bradley et al., 1990; Merutka et al., 1993).

Next to NOEs, $^3J_{HN\alpha}$ coupling constants are perhaps the most useful parameters for determining the local conformation (Wüthrich, 1986). Through the Karplus relationship, $^3J_{HN\alpha}$ coupling constants can be related directly to the ϕ backbone dihedral angle (Karplus, 1959; Wüthrich, 1986; Case et al., 1994). The angle ϕ takes on distinct values for a helix and an extended β-strand and serves as a fundamental discriminator of local secondary structure. There are several methods of determining $^3J_{HN\alpha}$ values in non-isotopically labeled peptides, most of which rely on variants of the [^1H,^1H] COSY experiment (Wüthrich, 1986). In an alanine-rich helical peptide, there is little dispersion of the $C^{\alpha}H$ resonances and this can lead to significant overlap in the $C^{\alpha}H$-NH spectral region, thereby obscuring the desired COSY signals.

*To whom correspondence should be addressed.
**Deceased March 5, 1996.

Abbreviations: $^3J_{HN\alpha}$, three-bond $C^{\alpha}H$-NH scalar coupling constant; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser spectroscopy; COSY, two-dimensional correlated spectroscopy; DQF-COSY, two-dimensional double-quantum-filtered correlated spectroscopy; TOCSY, two-dimensional total correlation spectroscopy.

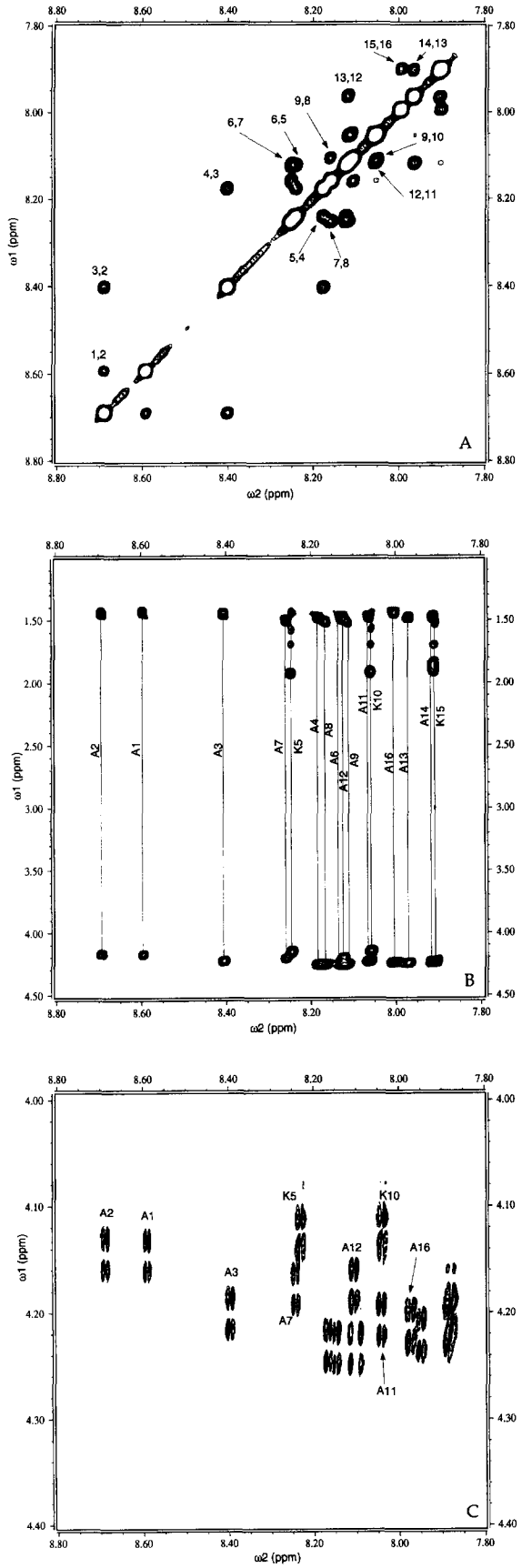

Fig. 1. 500 MHz ^{1}H NMR spectra of the 3K peptide. (A) NH-NH region of the NOESY spectrum with $\tau_m = 400$ ms; (B) TOCSY spectrum with $\tau_m = 60$ ms; (C) DQF-COSY spectrum.

Recently, Szyperski and co-workers have developed a novel method for extracting vicinal coupling constants of spins with a single-coupling partner (Szyperski et al., 1992). Their approach uses an inverse Fourier transform and iterative fitting in the time domain to obtain the line shape of individual NOEs. For a well-resolved NOE at the ω_2 of a particular amino acid backbone NH, the line shape parallel to the ω_2 axis directly reflects both the uncoupled line-width and the $^3J_{HN\alpha}$ value. The line shape of a particular NH resonance is reflected in all observable NOEs at the ω_2 corresponding to that resonance, so that data may be extracted from several regions of the NOESY spectrum. This feature greatly increases the likelihood of measuring $^3J_{HN\alpha}$ for a particular residue.

The method of Szyperski et al. is applied here to the homonuclear ^{1}H NOESY spectrum of the alanine-rich sequence (Marqusee et al., 1989):

Ac-AAAAKAAAAKAAAAKA-NH_2 3K

where A = alanine and K = lysine, in order to measure the $C^{\alpha}H$-NH coupling constant for each residue. The $^3J_{HN\alpha}$ value is determined in duplicate for each of the sixteen residues. A DQF-COSY gives unobstructed antiphase doublets for nine of the 16 residues in the 3K peptide and the resulting values agree well with those determined by the method of Szyperski et al. A plot of the $^3J_{HN\alpha}$ values versus the sequence provides new insight into how the helix content is distributed within the 3K peptide.

The 3K peptide was synthesized and purified according to previously reported protocols (Miick et al., 1993). All NMR experiments were performed with 7–8 mM samples at pH 5.0 with 10% D_2O in standard 5-mm NMR tubes. Data were processed using the MNMR package (Carlsberg Laboratory, Department of Chemistry, Copenhagen, Denmark) and analyzed using XEASY (Bartels et al., 1995).

All NMR experiments were performed at 500 MHz. NOESY spectra were acquired with $n_1 = 1024$ and $n_2 = 1024$. Solvent saturation was applied for 2 s before the first 90° pulse as well as during the mixing time. Data were zero-filled to 2048×2048 and multiplied by a shifted ($\pi/2$ in t_2 and $\pi/2.4$ in t_1) sine-bell window (sine-squared in t_1) before Fourier transformation. The TOCSY spectrum was acquired with $\tau_M = 60$ ms, $n_1 = 256$ and $n_2 = 1024$. Solvent saturation was applied for 2 s before the first 90° pulse. Data were zero-filled to 512×2048 and multiplied by a shifted sine-squared window before Fourier transformation. The DQF-COSY spectrum was acquired with $n_1 = 1024$ and $n_2 = 4096$. Solvent saturation was applied for 2 s before the first 90° pulse. Data were zero-filled to 2048×8192 and, in the t_1 dimension (but not the t_2 dimension), multiplied by a sine window before Fourier transformation. The DQF-COSY antiphase doublets were fitted to two Lorentzian functions and $^3J_{HN\alpha}$ coupling

constants were determined by the difference in the center frequencies. The resulting values were confirmed using the method of Kim and Prestegard (Kim and Prestegard, 1989).

Three NOESY spectra ($\tau_m = 100$, 200 and 400 ms), a TOCSY spectrum and a DQF-COSY spectrum were obtained from an approximately 7-mM solution of a 3K peptide at pH 5.0 in phosphate buffer at 2 °C (Fig. 1). In helical peptides, there is typically a good dispersion of the backbone NH protons. Relative to random coil values, the N-terminal NH protons shift downfield and the C-terminal NH protons shift upfield. Thus, the amide region is well-resolved and NOESY spectra are assignable using both the NH(i) – NH(i+1) NOE cross-peaks and the side-chain spin systems identified in the TOCSY. Subsequently, the $\tau_m = 200$ ms NOESY spectrum was processed along ω_2 according to the INFIT procedure of Szyperski et al. (as incorporated into the program XEASY (Bartels et al., 1995)). For each residue, $^3J_{HN\alpha}$ was measured at two locations: the diagonal NH(i,i) peak and one selected NOE cross-peak, usually the $C^\beta H(i) - NH(i)$. While the $C^\beta H(i) - NH(i)$ cross-peaks are generally well resolved, several of the NH(i,i) peaks are partially overlapped and for these residues determination of $^3J_{HN\alpha}$ solely from the diagonal is not unambiguous. In order to obtain $^3J_{HN\alpha}$ estimates from the NH(i,i) peaks in these cases, slices were taken where a single NH proton dominated the line shape. For example, the NH(5,5) and NH(7,7) peaks partially overlap. $^3J_{HN\alpha}(5)$ was determined from the upfield shoulder of this overlapped peak and $^3J_{HN\alpha}(7)$ was determined from the downfield shoulder. The diagonal and off-diagonal determinations are plotted against each other in Fig. 2. There is a good agreement between the two determinations, with no observable systematic errors arising from the region in which the $^3J_{HN\alpha}$ values were measured. The largest deviation from the mean corresponds to an error of less than 0.3 Hz.

Figure 3 shows the average of the two NOESY determinations of $^3J_{HN\alpha}$ plotted against the peptide sequence. Also shown are the DQF-COSY results for the nine well-resolved residues. (The remaining DQF-COSY signals overlap.) The agreement between the two methods is excellent. A recent comparison of the crystallographically determined secondary structure to the measured coupling constants suggests that $^3J_{HN\alpha}$ is approximately 5.0 Hz for helices (or more specifically 4.8 Hz for α-helices and 5.6 Hz for 3_{10}-helices) and 8.5 Hz for extended β-strands (Smith et al., 1996). Studies of proteins suggest that three sequential $^3J_{HN\alpha}$ values of less than 6.0 Hz are a good indication of a helical turn (Wüthrich, 1986). All of the $^3J_{HN\alpha}$ values, except for Lys^{15}, are less than 6.0 Hz, further confirming that the 3K sequence is mostly helical. From Ala^4 to Ala^{13} the $^3J_{HN\alpha}$ value ranges from 4.5 to 5.0 Hz, indicating a regular helix structure. Assuming that ϕ is bounded within the range $-120° < \phi < -25°$, ϕ is readily determined from $^3J_{HN\alpha}$ (using the paramaters of Vuister and Bax (1993)) and the results are also shown in Fig. 3.

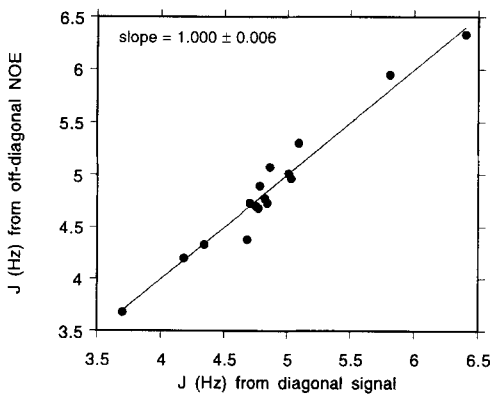

Fig. 2. $^3J_{HN\alpha}$ constants determined for each residue in the 3K peptide according to the method of Szyperski et al. (1992). Diagonal signals were obtained from the NH(i,i) peaks in the 2D NOESY spectrum. Most of the off-diagonal NOEs were determined from $C^\beta H$-NH cross-peaks. The line represents a least-squares fit to $y = m\,x$ with slope and error as indicated in the figure.

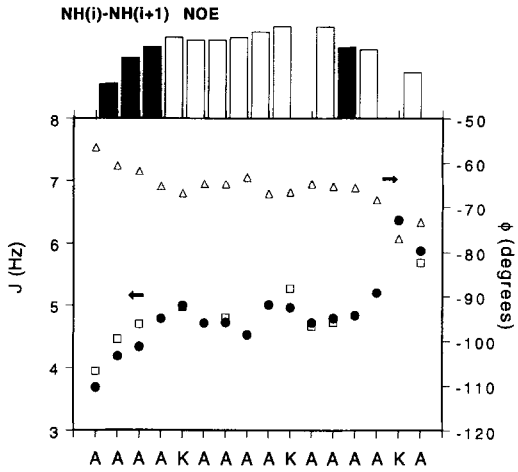

Fig. 3. Resulting $^3J_{HN\alpha}$ values (●, □; left axis) and corresponding ϕ dihedral angles (Δ; right axis) for the 3K sequence at 2 °C. $^3J_{HN\alpha}$ values were determined by both the inverse Fourier transform procedure (●) and DQF-COSY (□); ϕ angles were determined from the inverse Fourier transform results. Sequential NH-NH NOEs are shown at the top of the figure. Unfilled bars indicate estimates from partially overlapping peaks.

In addition to overall helicity, the results in Fig. 3 exhibit several interesting trends. First, as mentioned before, the C-terminal residues have the largest $^3J_{HN\alpha}$ (and the largest negative ϕ) values. The measured values for $^3J_{HN\alpha}$ reflect an ensemble average and Fig. 3 suggests that there is a significant population of non-helical conformers contributing to the averaged structure at the C-terminus. Fraying at the C-terminus has been observed in studies on this peptide and similar helical sequences (Chakrabartty et al., 1991; Liff et al., 1991; Miick et al., 1993; Rohl and Baldwin, 1994). Furthermore, the lysine residues, as well as Ala^4 and Ala^9, each of which precedes a lysine, all show slightly increased values for $^3J_{HN\alpha}$ compared to neighboring residues. This subtle trend may

indicate a local distortion and/or a slightly higher tendency for the helix to unfold at lysine residues. The helix propensity of lysine is presumably lower than that for alanine (O'Neil and De Grado, 1990) and increasing the lysine content in 3K-like sequences is helix-destabilizing (Marqusee et al., 1989). Following both Lys^5 and Lys^{10} are stretches of three sequential alanine residues with small $^3J_{HN\alpha}$ values, indicating regions of local helical stability.

The three N-terminal residues give rather small coupling constants, although $^3J_{HN\alpha}$ values within the range of 3–5 Hz are not unexpected for a helix (Smith et al., 1996). One possible interpretation of these small $^3J_{HN\alpha}$ values is that the helical structure persists through the N-terminus of the peptide. Experimental and theoretical studies reveal differences between helix termini, with the N-terminus often being the more stable one (Liff et al., 1991; Tirado-Rives and Jorgensen, 1991; Tobias and Brooks, 1991; Miick et al., 1993; Doig et al., 1994). What is rather perplexing, however, is that the N-terminus yields $^3J_{HN\alpha}$ values smaller in magnitude than those for the central region of the helix. While the 3K N-terminus may be more stable than the C-terminus, there is still substantial evidence suggesting that the N-terminus is frayed relative to the middle of the peptide. NH(i) – NH(i+1) NOEs are diagnostic for a helix and the volume integrals from the 3K spectrum with $\tau_m = 200$ ms are shown in Fig. 3. A reduced NOE intensity is evident at both helix termini. Thus, both ends of the peptide are frayed, but it appears that the mechanism of unfolding is different for the respective termini. Unfolding at the C-terminus clearly requires motion of both ϕ and Ψ, as indicated by the loss of NH(i) – NH(i+1) contact, concomitant with an increase in the $^3J_{HN\alpha}$ value. However, local unfolding of N-terminal residues approximately preserves the helical values of $^3J_{HN\alpha}$ and thus may involve excursions mainly along the Ψ coordinate, with ϕ retaining a helical value. Recent molecular dynamics calculations have indeed suggested that helix propagation (the opposite of local unfolding) follows unique pathways at the respective helix termini (W.S. Young and C.L. Brooks, III (1996) *J. Mol. Biol.*, in press). Associated with this phenomenon is the suggestion that the amino acid helix propensities differ at the helix ends. N-terminal propagation is more cooperative and favoured over C-terminal propagation. The experiments presented here reveal distinct fraying mechanisms for the helix termini and lend support to the observations of Brooks and co-workers.

The results presented here demonstrate that complete profiles of $^3J_{HN\alpha}$ coupling constants are obtainable from peptides with highly repetitive alanine-rich sequences. The method is straightforward, reliable and does not require isotopic substitution or modification of the alanine-rich design. When compared to other measures of local helix content, the $^3J_{HN\alpha}$ values obtained provide new insights into the fraying mechanism at the helix termini. This approach should prove to be generally applicable for probing position-dependent helicity and exploring the subtle factors that influence helix propagation and structure.

Acknowledgements

The authors thank Dr. Siggi Mronga for technical help with the INFIT procedure. The software packages MNMR and XEASY were kindly provided by M. Kjaer and K. Wüthrich, respectively. This work was supported by an NSF grant MCB 9408284 and NIH grants GM46870 (G.L.M.) and GM16716 (C.J.S.). Much of this work was performed while G.L.M. was on sabbatical leave at the University of Nijmegen. G.L.M. wishes to extend his sincere gratitute to Prof. C. Hilbers and the members of the Biophysical Laboratory for their help and hospitality.

References

Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **5**, 1–10.

Bradley, E.K., Thomason, J.F., Cohen, F.E., Kosen, P.A. and Kuntz, I.D. (1990) *J. Mol. Biol.*, **215**, 607–622.

Case, D.A., Dyson, H.J. and Wright, P.E. (1994) *Methods Enzymol.*, **239**, 392–416.

Chakrabartty, A., Schellman, J.A. and Baldwin, R.L. (1991) *Nature*, **351**, 586–588.

Doig, A.J., Chakrabartty, A., Klingler, T.M. and Baldwin, R.L. (1994) *Biochemistry*, **33**, 3396–3403.

Fiori, W.R., Miick, S.M. and Millhauser, G.L. (1993) *Biochemistry*, **32**, 11957–11962.

Karplus, M. (1959) *J. Chem. Phys.*, **30**, 11–15.

Kim, Y.M. and Prestegard, J.H. (1989) *J. Magn. Reson.*, **84**, 9–13.

Liff, M.I., Lyu, P.C. and Kallenbach, N.R. (1991) *J. Am. Chem. Soc.*, **113**, 1014–1019.

Lyu, P.C., Liff, M.I., Marky, L.A. and Kallenbach, N.R. (1990) *Science*, **250**, 669–673.

Marqusee, S., Robbins, V.H. and Baldwin, R.L. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5286–5290.

Merutka, G., Morikis, D., Brüschweiler, R. and Wright, P.E. (1993) *Biochemistry*, **32**, 13089–13097.

Merutka, G. and Stellwagen, E. (1990) *Biochemistry*, **29**, 894–898.

Miick, S.M., Casteel, K.M. and Millhauser, G.L. (1993) *Biochemistry*, **32**, 8014–8021.

O'Neil, K.T. and De Grado, W.F. (1990) *Science*, **250**, 646–651.

Osterhout Jr., J., Baldwin, R.L., York, E.J., Stewart, J.M., Dyson, H.J. and Wright, P.E. (1989) *Biochemistry*, **28**, 7059–7064.

Rohl, C.A. and Baldwin, R.L. (1994) *Biochemistry*, **33**, 7760–7767.

Smith, L.J., Bolin, K.A., Schwalbe, H., Macarthur, M.W., Thornton, J.M. and Dobson, C.M. (1996) *J. Mol. Biol.*, **255**, 494–506.

Storrs, R.W., Truckses, D. and Wemmer, D.E. (1992) *Biopolymers*, **32**, 1695–1702.

Szyperski, T., Güntert, P., Otting, G. and Wüthrich, K. (1992) *J. Magn. Reson.*, **99**, 552–560.

Tirado-Rives, J. and Jorgensen, W.L. (1991) *Biochemistry*, **30**, 3864–3871.

Tobias, D.J. and Brooks, C.L. (1991) *Biochemistry*, **30**, 6059–6070.

Vuister, G.W. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 7772–7777.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.

Zhou, H.X.X., Hull, L.A., Kallenbach, N.R., Mayne, L., Bai, Y.W. and Englander, S.W. (1994) *J. Am. Chem. Soc.*, **116**, 6482–6483.