Sequence-Corrected ^{15}N "Random Coil" Chemical Shifts

Daniel Braun, Gerhard Wider, and Kurt Wüthrich*

Contribution from the Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland

Received April 11, 1994[®]

Abstract: The ^{15}N chemical shifts in a group of oligopeptides H-Gly 1-Gly 2-Xxx 3-Ala 4-OH, where Xxx stands for one of the 20 proteinogenic amino acid residues, were measured using two-dimensional heteronuclear single-quantum coherence (HSQC) experiments. Following the hypothesis, based on earlier observations, that the "random coil" ^{15}N shifts are predominantly determined by the chemical structures of the residue considered and the preceding residue in the sequence, we evaluated the influence of the residue type in position 3 on the ^{15}N chemical shift of Ala 4 in the above peptides. These data were combined with the ^{15}N chemical shifts for the 20 common amino acid residues measured in position 3 of the same peptides to calculate the 400 "random coil" reference shifts for the 20 amino acid residues Xxx in all possible dipeptide sequences Yyy–Xxx. Chemical shift predictions based on this data set fit experimental ^{15}N shifts in denatured proteins within a range of ±2 ppm and, thus, represent a useful reference for investigations on correlations between ^{15}N chemical shifts and polypeptide conformation.

Introduction

For NMR[1] spectroscopists working with proteins, the ^{15}N chemical shifts of the backbone amide groups in polypeptide chains are of considerable practical interest since they cover a wide range of about 40 ppm. For example, in the "fingerprint region" containing the cross peaks of a backbone atom group of each residue,[2] the peaks are usually much better separated in a 2D [^{15}N,^{1}H]-COSY[3] spectrum than in a homonuclear 2D [^{1}H,^{1}H]-COSY[2] spectrum, and the ^{15}N chemical shift dispersion along the third frequency axis generally results in dramatically improved spectral resolution in a 3D ^{15}N-resolved [^{1}H,^{1}H]-correlation experiment when compared to the corresponding 2D [^{1}H,^{1}H]-experiment.[4] Somewhat in contrast with this obvious practical interest, attempts to correlate ^{15}N chemical shifts with specific features of three-dimensional protein structures have met with only limited success.[5] This is due in part to the fundamental problem that a counterpart to the "random coil" ^{1}H chemical shifts in proteins[2,6] is difficult to define for ^{15}N, since backbone peptide ^{15}N chemical shifts for a given residue type are strongly affected by the neighboring residue types in the sequence.[5] This is best illustrated by the fact that the dispersion of the ^{15}N shifts of an unfolded protein in 7 M urea solution was sufficient to enable complete resolution and assignment of the NMR spectrum.[7] The present paper introduces a complete set of sequence-corrected ^{15}N chemical shifts derived from experiments with model peptides, which provide a reference for future studies of correlations between ^{15}N shifts and protein conformation.

On fundamental grounds[8] the dominant sequence effects on the ^{15}N chemical shift of a given residue type arise from the nature of the sequentially preceding residue. This fact has previously been used as a basis for sequence corrections to "random coil" ^{15}N shifts.[5] However, for the calibration of these corrections the authors used data from a variety of short peptides and protected amino acids measured in different solvents. The application of the resulting "random coil" values for the interpretation of ^{15}N chemical shifts of unfolded proteins still shows rather large discrepancies (see below). The ^{15}N shifts for the 20 common amino acid residues presented in this paper have all been derived from a set of homologous oligopeptides for all 20 common amino acid residues. The resulting data set provides for much improved fits of the ^{15}N NMR spectra of unfolded polypeptide chains.

Experimental Section

Nineteen tetrapeptides of the type H-Gly-Gly-Xxx-Ala-OH and the pentapeptide H-Gly-Gly-Lys-Ala-Ala-OH were purchased from Bachem AG, Liestal, Switzerland. For the NMR experiments the peptides were dissolved in 90% H_2O/10% $^{2}H_2O$ at a concentration of 100 mM. H-Gly-Gly-Cys-Ala-OH was prepared by reduction of the corresponding dimeric cystine peptide obtained from Bachem AG with 250 mM sodium dithionite. The measurements were performed at 35 °C and at the two pH values 2.0 and 5.0. At pH 5.0, 20 mM acetic acid was added as a buffer. The two pH values have been chosen to be well above and well below the pK_a of the C-terminal carboxyl group (p$K_a \approx 3.5$), respectively, and to be close to the minimum exchange rate for amide protons, which is near pH 3.0.[2] To measure the influence of the side chain protonation of Asp and Glu, the protected tetrapeptides TFA-Gly-Gly-Asp-Ala-OCH_3 and TFA-Gly-Gly-Glu-Ala-OCH_3 were titrated over the pH range 2.0–6.5, whereby 100 mM NaCl was added to minimize the relative change of the ionic strength during the titration.

The NMR spectra were recorded on a Bruker AM360 NMR instrument, using the two-dimensional HSQC experiment[9] at the natural abundance of ^{15}N. The sweep width in the ^{15}N dimension was 15.22 ppm. The data were acquired with the TPPI-states method, using a first delay of t_1 that was half as long as the increment, which ensures opposite sign for folded and unfolded resonances.[10] t_{1max} was 230 ms, t_{2max} was 256 ms, and the time domain data size was 256 × 2048 points. The total measuring time was about 2 h per experiment. The data matrices were zero-filled 4-fold along t_1, resulting in a digital resolution of 0.03 ppm in the ^{15}N dimension. The ^{1}H chemical shifts were referenced to [2,2,3,3-$^{2}H_4$]-(trimethylsilyl)propionate (TSP), and the ^{15}N chemical shifts were referenced indirectly to external liquid NH_3.[11]

The ^{15}N chemical shift of Pro was measured in a 1D ^{15}N spectrum using a 500 mM solution of the tetrapeptide H-Gly-Gly-Pro-Ala-OH in

® Abstract published in *Advance ACS Abstracts,* August 15, 1994.

(1) Abbreviations used: NMR, nuclear magnetic resonance; ppm, parts per million; 1D, one dimensional; 2D, two dimensional; 3D, three dimensional; COSY, 2D correlation spectroscopy; HSQC, heteronuclear single-quantum coherence; TPPI, time-proportional phase incrementation; TFA, protection group: CF_3CO–.

(2) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(3) Wüthrich, K.; Spitzfaden, C.; Memmert, K.; Widmer, H.; Wider, G. *FEBS Lett.* **1991**, *285*, 237–247.
(4) Marion, D.; Driscoll, P. C.; Kay, L. E.; Wingfield, P. T.; Bax, A.; Gronenborn, A.M.; Clore, G. M. *Biochemistry* **1989**, *28*, 6150–6156.
(5) (a) Glushka, J.; Lee, M.; Coffin, S.; Cowburn, D. *J. Am. Chem. Soc.* **1989**, *111*, 7716–7722; (b) *J. Am. Chem. Soc.* **1990**, *112*, 2843.
(6) Wüthrich, K. *NMR in Biological Research: Peptides and Proteins*; North-Holland: Amsterdam, 1976.
(7) Neri, D., Wider; G., Wüthrich, K. *Proc. Natl. Acad. Sci.* U.S.A. **1992**, *89*, 4397–4401.
(8) Martin, G. J.; Martin M. L.; Gouesnard, J. *^{15}N-NMR Spectroscopy*; Springer: New York, 1981.
(9) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185–189.
(10) Bax, A.; Marion, D. *J. Magn. Reson.* **1988**, *78*, 186–196.
(11) Live, D. H.; Davis, D. G.; Agosta, W. C.; Cowburn, D. *J. Am. Chem. Soc.* **1984**, *106*, 1939–1941.

0002-7863/94/1516-8466$04.50/0 © 1994 American Chemical Society

Table 1. ^{15}N Chemical Shifts, $\delta(^{15}N)$, of the Tetrapeptides H-Gly 1-Gly 2-Xxx 3-Ala 4-OH in H_2O Solution at 35 °C

	$\delta(^{15}N)$,[a] pH = 2.0			$\delta(^{15}N)$,[a] pH = 5.0		
peptide	Gly2	Xxx 3	Ala 4	Gly2[b]	Xxx 3	Ala 4
GGGA	108.55	108.72	124.32	108.68	109.30	129.03
GGAA	108.46	123.75	124.16		124.40	128.66
GGVA	108.49	119.14	128.54		119.17	132.51
GGIA	108.43	119.82	128.82		119.99	132.82
GGLA	108.37	121.85	125.86	108.48	122.24	129.71
GGSA	108.48	115.62	126.31		116.05	130.96
GGTA	108.43	113.93	127.52	108.50	113.70	131.55
GGDA	108.52	118.84	125.08	108.65	120.45	128.82
GGEA	108.35	119.77	126.32	108.75	120.83	129.88
GGNA	108.35	118.53	125.05		118.99	129.26
GGQA	108.40	119.91	126.43	108.54	120.39	130.99
GGMA	108.32	119.92	126.46		120.33	130.44
GGCA	108.38	118.11	126.33	108.40	118.54	130.48
(GGCA)$_2$	108.21	118.45	126.62	108.34	118.89	130.81
GGWA	108.23	120.68	126.59		120.81	130.26
GGFA	108.28	119.91	126.62		119.98	130.24
GGYA	108.30	119.99	126.65	108.30	120.04	130.26
GGHA[c]	108.20	118.10	127.12		118.00	131.06
GGPA[d]	108.59	133.87	124.89			129.56
GGRA	108.46	120.69	126.61		121.06	131.00
GGKAA[e]	108.45	121.09	125.81		121.15	126.41

[a] The chemical shifts are in ppm ± 0.05 ppm relative to external NH_3 as described in ref 11. Dr. D. Live (personal communication) recently informed us that while this calibration used a sample of CH_3NO_2 and TMS with $CDCl_3$ added as a lock substance, a new calibration using 90% CH_3NO_2 in d_6-acetone and a small amount of TMS gives a somewhat different reference value. When using this new calibration, all the values in this table and in Table 2 would have to be increased by 1.8 ppm. [b] At pH 5.0 the chemical shift of Gly 2 could not be observed in all cases, because of rapid exchange of this amide proton with H_2O. [c] The chemical shifts of GGHA at pH 10.0 are 108.29 ppm for Gly 2, 119.41 ppm for His 3, and 130.47 ppm for Ala 4. The chemical shifts of the reference GGGA at pH 10.0 are 108.61 ppm and 109.17 ppm for Gly 2 and Gly 3 (not uniquely identified) and 129.22 ppm for Ala 4. [d] Only the form containing the *trans* Gly–Pro peptide bond could be observed due to limited signal to noise ratio in the spectra. [e] The chemical shift of Ala 5 is 124.40 ppm at pH = 2.0 and 128.74 ppm at pH = 5.0.

90% H2O/10% ^2H_2O. During data acquisition, the protons were decoupled applying WALTZ with a field strength of 1000 Hz. A total of 100 000 scans with 3 s delay time have been accumulated. In the same way the ^{15}N resonances of 200 mM samples of H-Gly-Gly-Gly-Ala-OH and H-Gly-Gly-His-Ala-OH were detected at pH 10.0, where the amide proton exchange is so fast that HSQC experiments cannot be used.

Results

^{15}N Resonance Assignments. At pH 2.0 the ^{15}N HSQC-spectra of the linear peptides H-Gly 1-Gly 2-Xxx 3-Ala 4-OH contained typically three peaks, because the exchange of the amino protons of Gly 1 was too fast for a separate signal to be detected. The sequence-specific assignment of the peaks was made primarily by comparing the ^{1}H chemical shifts with those of Bundi and Wüthrich.[12,13] Independently, the resonance of Gly 2 could readily be identified in the 1D ^{1}H NMR spectrum by the fact that it is the only triplet in the low-field region. Going from pH 2.0 to 5.0, the deprotonation of the C-terminal carboxyl group gives rise to a low-field shift of the amide ^{15}N-resonance of Ala 4 of approximately 4 ppm, whereas the residue in position 3 is nearly unaffected. The resonance of Ala 4 in H-Gly-Gly-Ala-Ala-OH could thus be identified by following the signal in a titration from pH 5.0 to 2.0. In about half of the tetrapeptides at pH 5.0 the amide proton of residue 2 exchanged too rapidly to be observed as a separate line.[12] The ^{15}N chemical shifts are listed in Table 1.

Calibration of the Parameters for Sequence-Corrections of ^{15}N Chemical Shifts. The ^{15}N chemical shifts of Gly 2 in all 20 tetrapeptides vary only over a range of 0.4 ppm (Table 1). This shows that sequence effects are dominated by the influence of the directly preceding residue, so that the previously proposed eq 1[5] is indeed a good approximation for the evaluation of the sequence-corrected "random coil" chemical shifts.

(12) Bundi, A.; Wüthrich, K. *Biopolymers* **1979**, *18*, 285–297.
(13) Bundi, A.; Wüthrich, K. *Biopolymers* **1979**, *18*, 299–311.

Table 2. Parameters δ_{rc}^{GX} and $\Delta\delta^{Y}$ To Be Used with eq 1 for Calculating Sequence-Corrected ^{15}N "Random Coil" Chemical Shifts, δ_{rc}^{YX}, for the 20 Common Amino Acid Residues

residue	δ_{rc}^{GX} [a] (ppm)	$\Delta\delta^{Y}$ (ppm)
Gly	108.7	0.0
Ala	123.7	−0.2
Val	119.1	4.2
Ile	119.8	4.5
Leu	121.8	1.5
Ser	115.6	2.0
Thr	113.9	3.2
Asp (pH < 2.5)[b]	118.8	0.8
Asp (pH > 5.5)[b]	120.4	0.3
Glu (pH < 3.0)[c]	119.8	2.0
Glu (pH > 6.0)[c]	120.8	1.8
Asn	118.5	0.7
Gln	119.9	2.1
Met	119.9	2.1
Cys	118.1	2.0
Trp	120.7	2.3
Phe	119.9	2.3
Tyr	120.0	2.3
His (pH < 5.5)[d]	118.1	2.8
His (pH > 8.5)[d]	119.4	1.2
Pro (*trans*)	133.9	0.6
Arg	120.7	2.3
Lys	121.1	2.1

[a] Chemical Shifts, δ_{rc}^{GX}, are relative to external liquid NH_3 (see Table 1). [b] pK_a of Asp in H-Gly-Gly-Asp-Ala-OH: 3.9.[12] [c] pK_a of Glu in H-Gly-Gly-Glu-Ala-OH: 4.3.[12] [d] pK_a of His in H-Gly-Gly-His-Ala-OH: 7.0.[12]

$$\delta_{rc}^{YX} = \delta_{rc}^{GX} + \Delta\delta^{Y} \quad (1)$$

δ_{rc}^{YX} is the sequence-corrected "random coil" chemical shift for residue Xxx preceded by residue Yyy, δ_{rc}^{GX} is the "random coil" chemical shift of residue Xxx preceded by Gly, and $\Delta\delta^{Y}$ is the increment accounting for the influence of replacing Gly by Yyy in the position preceding the residue considered.

For use with the experimental data of Table 1, eq 1 is calibrated as follows: δ_{rc}^{GX} is the chemical shift for residue Xxx 3 in the peptides H-Gly-Gly-Xxx-Ala-OH, and $\Delta\delta^{Y}$ is the chemical shift difference for residue Ala 4 between H-Gly-Gly-Yyy-Ala-OH and H-Gly-Gly-Gly-Ala-OH. On the basis of the following additional considerations, we propose to use the parameter set of Table 2 to evaluate sequence-corrected ^{15}N "random coil" chemical shifts for the 20 common amino acid residues. (i) Between pH 2.0 and 5.0, the largest change of the ^{15}N chemical shift of residue 3 is 0.7 ppm except when Asp or Glu is in position 3 (see below). These small shifts are probably due to conformational effects induced by the additional charge at pH 5.0. To exclude possible influences of such charge effects, the parameters in Table 2 were obtained from the measurements at pH 2.0. (ii) For Lys, where the pentapeptide H-Gly-Gly-Lys-Ala-Ala-OH was measured, $\Delta\delta^{K}$ was taken as the difference between the ^{15}N shifts of Ala 4 in H-Gly-Gly-Lys-Ala-Ala-OH and Ala 3 in H-Gly-Gly-Ala-Ala-OH. (iii) To assess the influence of side chain deprotonation, the protected tetrapeptides of TFA-Gly-Gly-Glu-Ala-OMe and TFA-Gly-Gly-Asp-Ala-OMe were measured over the pH range 2.0–6.5 (Figure 1). The titration shifts for the residues Xxx 3 coincide closely with the measurements of the unprotected tetrapeptides and show that the influence of pH cannot be neglected for these two side chains. Therefore, two sets of parameters for the protonated and deprotonated forms of Asp and Glu are included in Table 2. The $\Delta\delta^{E}$ and $\Delta\delta^{D}$ values of the deprotonated form were calculated adding the titration shifts of Ala 4 (Figure 1) to their values at pH 2.0. (iv) To

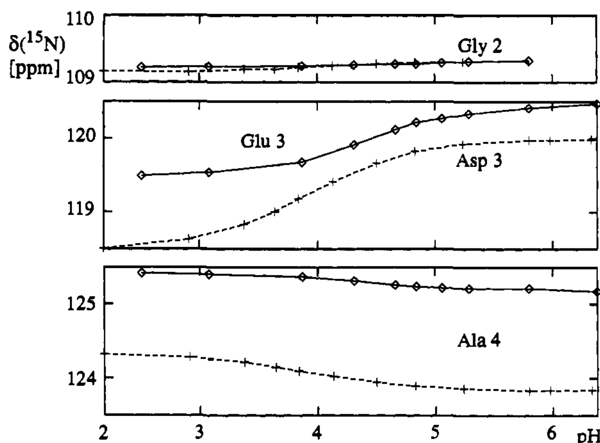

Figure 1. pH dependence of the ^{15}N chemical shifts of TFA-Gly-Gly-Glu-Ala-OMe (solid lines) and TFA-Gly-Gly-Asp-Ala-OMe (dashed lines). Peptide concentration 100 mM in 90% H2O/10% $^{2}H_2O$, 100 mM NaCl, T = 35 °C.

account for the influence of the titration of His, a measurement of H-Gly-Gly-His-Ala-OH at pH 10 was performed. δ_{rc}^{GH} and $\Delta\delta^{H}$ for the deprotonated form were calculated using the chemical shifts of H-Gly-Gly-Gly-Ala-OH at pH 10 as a reference. (v) Since closely similar ^{15}N shifts were obtained for the cysteinyl tetrapeptide and the cystine dimer (Table 1), only one parameter set for Cys is included in Table 2.

Comparison of ^{15}N Chemical Shift Predictions with Measurements in Unfolded Polypeptide Chains. Sequence-specific NMR assignments for unfolded proteins are still scarce,[14] and chemical shift lists have not been presented for all proteins for which resonance assignments were described.[15] Two data sets available to us have been used to prepare assessments of the quality of ^{15}N "random coil" chemical shift prediction with the presently proposed approach of eq 1 and the data base of Table 2.

The protein GAL 4 forms a well-defined globular core of residues 10–40, with flexibly disordered chain ends of residues 1–9 and 41–62.[16,17] Complete sequence-specific ^{1}H, ^{13}C, and ^{15}N resonance assignments for the polypeptide backbone have been obtained,[17] and the differences between the chemical shifts for this protein and the "random coil" shifts have been plotted in Figure 2. It has long been established that ^{1}H chemical shifts are sensitive probes of three-dimensional polypeptide folding,[2,6] and recently correlations between $^{13}C^{\alpha}$ shifts and secondary structure have been established.[19] The plots for the amide protons, the α protons, and $^{13}C^{\alpha}$ clearly manifest the aforementioned globular core from residues 10–40 of GAL 4 by large deviations from the "random coil" shifts. The "random coil" nature of the chain terminal polypeptide segments 1–9 and 41–62 is equally well reflected in these data, since the only significant deviations from "random coil" shifts are seen in the two dipeptide segments -Ser 41-Pro 42- and -Ser 47-Pro 48-. The plot for $\Delta\delta(^{15}N)$ confirms that with the presently proposed approach for estimating ^{15}N chemical shifts, qualitatively similar information can be obtained from $\delta(^{15}N)$ as from the ^{1}H and $^{13}C^{\alpha}$ shifts. In particular, for the flexible chain ends the ^{15}N shifts are near the "random coil" values, again with the sole exceptions of -Ser 41-Pro 42- and -Ser 47-Pro 48- (arrows in Figure 2).

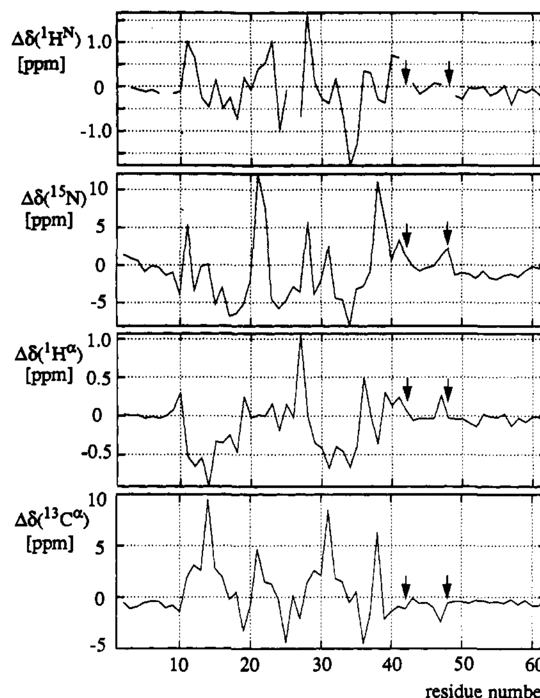

Figure 2. Plots *versus* the amino acid sequence of the protein GAL 4 of the conformation-dependent chemical shifts of the amide protons, amide ^{15}N, α protons, and α carbons. The values plotted correspond to the difference between the experimental shifts for GAL 4[16] and the "random coil" shifts for protons,[12] carbon-13,[18] and nitrogen-15 (eq 1 and Table 2). The arrows identify the positions of Pro 41 and Pro 47 (see text).

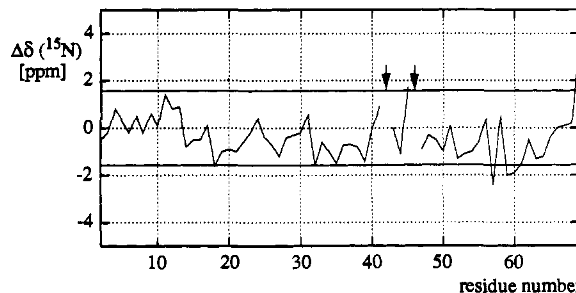

Figure 3. Plot *versus* the amino acid sequence of the differences between the ^{15}N chemical shifts measured for the unfolded DNA-binding domain of the 434 repressor in 7 M urea at pH 4.8 and the "random coil" reference shifts predicted using eq 1 and Table 2. The values for Glu and Asp were interpolated (Table 2). The arrows identify the positions of Pro 42 and Pro 46 for which no ^{15}N chemical shifts are available. The two solid lines indicate the chemical shift range of ±1.6 ppm (see text). The experimental shifts from the 434 repressor were recalibrated to the presently used external liquid NH_3 reference by the addition of 1.3 ppm to the values reported in ref 7.

For the DNA-binding domain of the 434 repressor, complete ^{1}H and ^{15}N resonance assignments were obtained for the unfolded form in 7 M urea.[7,20] Figure 3 shows that the measured chemical shifts coincide within ±1.6 ppm with the "random coil" shifts predicted using eq 1 and Table 2, except for some residues near the polypeptide segment 54–60, for which residual non-random structure has been observed[21] and Leu 45 which is again a residue preceding a Pro.

For a comparison of the present approach of eq 1 and Table 2 with the previously proposed set of ^{15}N "random coil" chemical shifts,[5] we evaluated RMSD values for the measured shifts in unfolded 434 repressor (Figure 3) and in the unstructured segments of GAL 4 (Figure 2) relative to the two reference data sets (Table 3). To obtain a somewhat broader sample, we further

(14) Wüthrich, K. *Curr. Opin. Struct. Biol.* **1994**, *4*, 93–99.
(15) Logan, T. M.; Olejniczak, E. T.; Xu, R. X.; Fesik, S. W. *J. Biomol. NMR* **1993**, *3*, 225–231.
(16) Kraulis, P. J.; Raine, A. R. C.; Gadhavi, P. L.; Laue, E. D. *Nature* **1992**, *356*, 448–450. Baleja, J. D.; Marmorstein R.; Harrison, S. C.; Wagner, G. *Nature* **1992**, *356*, 450–453.
(17) Shirakawa, M.; Fairbrother, W. J.; Serikawa, Y.; Ohkubo, T.; Kyogoku, Y.; Wright, P. E. *Biochemistry* **1993**, *32*, 2144–2153.
(18) Richarz, R.; Wüthrich, K. *Biopolymers* **1978**, *17*, 2133–2141.
(19) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311–333. Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490–5491.
(20) Neri, D.; Wider, G.; Wüthrich, K. *FEBS Lett.* **1992**, *303*, 129–135.
(21) Neri, D.; Billeter, M.; Wider, G.; Wüthrich, K. *Science* **1992**, *257*, 1559–1563.

Table 3. RMSDs between ^{15}N Chemical Shifts Measured in "Unstructured" Polypeptide Chains and the Sequence-Corrected "Random Coil" Shifts Proposed in This Paper and Those Suggested Previously,[5b] Respectively

	RMSD (ppm)[b]	
protein[a]	δ_{rc}^{YX}	ref 5
434 (2–68)	0.8	1.7
GAL 4 (2–9, 43–61)	0.8	2.0
LFB1/HNF1 (2–14)	1.0	1.9
III^{Glc} (2–17)	1.1	2.3
IIA^{glc} (2–12)	0.8	1.3

[a] 434 stands for the urea-unfolded 434 repressor,[20] GAL 4 for the DNA-binding domain of GAL 4,[16] LFB1/HNF1 for the homeodomain of this transcription factor,[22] III^{Glc} for a signal-transducing phosphocarrier protein of *Escherichia coli*,[23] and IIA^{glc} for a domain of glucose permease from *Bacillus subtilis*.[24] For each protein the numbers in parentheses identify the polypeptide segment considered. [b] The two columns list the deviations from the values predicted using the presently proposed reference values δ_{rc}^{YX}, or the previously suggested values by Glushka *et al.*[5] To correct for imperfect chemical shift calibration, the RMSDs are calculated relative to the mean values of the deviations. Proline and the residues preceding Pro were not considered for the calculation.

included data on unstructured chain ends in proteins for which the necessary resonance assignments are available. These are a homeodomain polypeptide from the LFB1/HNF1 transcription factor,[22] the domain III^{Glc} of a signal-transducing phosphocarrier protein from *E. coli*,[23] and the glucose permease domain IIA^{glc} [24] from *Bacillus subtilis*. Table 3 shows that for the unstructured segments in all these proteins the RMSD values relative to the "random coil" ^{15}N shifts proposed in this paper are of the order of 1 ppm, whereas the corresponding RMSD values relative to the reference data of Glushka *et al.*[5] are of the order of 2 ppm.

(22) Leiting, B.; DeFrancesco, R.; Tomei, L.; Cortese, R.; Otting, G.; Wüthrich, K. *EMBO J.* **1993**, *12*, 1797–1803.
(23) Pelton, J. G.; Torchia, D. A.; Meadow, N. D.; Wong, C.; Roseman, S. *Biochemistry* **1991**, *30*, 10043–10057.
(24) Fairbrother, W. J.; Cavanagh, J.; Dyson, H. J.; Palmer, A. G.; Sutrina, S. L.; Reizer, J.; Saier, M. H.; Wright, P. E. *Biochemistry*, **1991**, *30*, 6896–6907. Fairbrother, W. J.; Palmer, A. G.; Rance, M.; Reizer, J.; Saier, M. H.; Wright, P. E. *Biochemistry* **1992**, *31*, 4413–4425.

Conclusions

The large dispersion of the backbone ^{15}N chemical shifts in proteins is due to three major factors, *i.e.*, the residue type, sequence-effects of the preceding residue of up to 5 ppm, and conformational effects of up to 15 ppm and more. The first two components can now be reliably predicted with the presently introduced sequence-corrected "random coil" chemical shifts, which provide a reference for further improved studies of the correlations between ^{15}N shifts and protein conformation.[25] Since the predicted ^{15}N "random coil" shifts coincide with experimental measurements only within about ±2 ppm (Figure 3), questions remain as to the causes of the residual discrepancies. Probably these are predominantly due to the previously noted fact that the tetrapeptides H-Gly-Gly-Xxx-Ala-OH adopt preferred conformations and are not really in a "random coil" state.[12,26] In particular, when comparisons are made with polypeptides in denaturing milieus (Figure 3), this may lead to systematic discrepancies. An obvious approach for further improving the ^{15}N chemical shift reference data set would be by direct measurement of the sequence effects for all combinations of residues, for example, in all 400 peptides H-Gly-Gly-Yyy-Xxx-Ala-OH. However, in view of the above it seems questionable whether this effort would be worthwhile, and the present approach with eq 1 and Table 2 could well become the standard for establishing ^{15}N "random coil" reference shifts for studies with proteins.

Acknowledgment. We thank Ch. Würsch for measuring part of the data, Mrs. E. Huber for the careful processing of the manuscript, and the Schweizerischer Nationalfonds for financial support (project 31.32033.91).

(25) de Dios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993**, *260*, 1491–1495.
(26) Wüthrich, K.; DeMarco, A, *Helv. Chim. Acta* **1976**, *59*, 2228–2235.