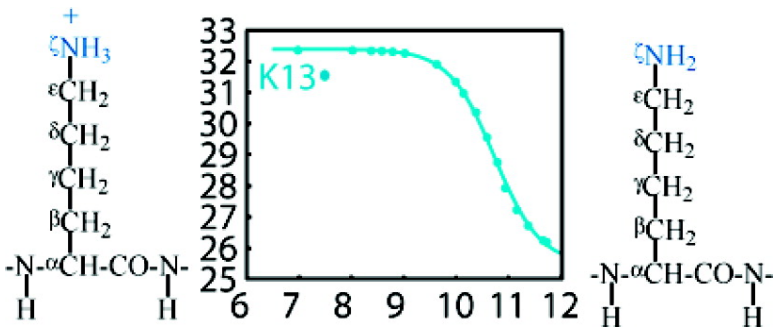
Article

Residue-Specific p*K* Determination of Lysine and Arginine Side Chains by Indirect N and C NMR Spectroscopy: Application to *apo* Calmodulin

Ingemar Andr, Sara Linse, and Frans A. A. Mulder

J. Am. Chem. Soc., **2007**, 129 (51), 15805-15813 • DOI: 10.1021/ja0721824



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 11/29/2007

Residue-Specific pK_a Determination of Lysine and Arginine Side Chains by Indirect ^{15}N and ^{13}C NMR Spectroscopy: Application to *apo* Calmodulin

Ingemar André,[†,§] Sara Linse,[†] and Frans A. A. Mulder*[,‡]

Contribution from the Department of Biophysical Chemistry, Lund University, P.O. Box 124, Sweden and Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4 9747 AG Groningen, the Netherlands

Received March 28, 2007; E-mail: f.a.a.mulder@rug.nl

Abstract: Electrostatic interactions in proteins can be probed experimentally through determination of residue-specific acidity constants. We describe here triple-resonance NMR techniques for direct determination of lysine and arginine side-chain protonation states in proteins. The experiments are based on detection of nonexchangeable protons over the full range of pH and temperature and therefore are well suited for pK_a determination of individual amino acid side chains. The experiments follow the side-chain $^{15}N^{\zeta}$ (lysine) and $^{15}N^{\epsilon}$ or $^{13}C^{\zeta}$ (arginine) chemical shift, which changes due to sizable changes in the heteronuclear electron distribution upon (de)protonation. Since heteronuclear chemical shifts are overwhelmed by the charge state of the amino acid side chain itself, these methods supersede ^{1}H-based NMR in terms of accuracy, sensitivity, and selectivity. Moreover, the $^{15}N^{\zeta}$ and $^{15}N^{\epsilon}$ nuclei may be used to probe changes in the local electrostatic environment. Applications to three proteins are described: *apo* calmodulin, calbindin D_{9k}, and FKBP12. For *apo* calmodulin, residue-specific pK_a values of lysine side chains were determined to fall between 10.7 and 11.2 as a result of the high net negative charge on the protein surface. Ideal two-state titration behavior observed for all lysines indicates the absence of significant direct charge interactions between the basic residues. These results are compared with earlier studies based on chemical modification.

Introduction

Electrostatic interactions are crucial in biology, providing the basis for long-range inter- and intramolecular forces. Specific interactions may modulate protein stability, protein−protein and protein−DNA/RNA interactions, enzyme catalysis, and conformational transitions, while the presence of charges on protein surfaces is necessary to maintain protein solubility and prevent aggregation. For example, the stability of thermophilic versus mesophilic cold shock proteins is due to surface charge interactions,[1] and this insight offers new routes to protein engineering.[2]

Positive charges in proteins are provided by histidine, arginine, and lysine residues and by the unprotected N-terminus. For example, RNA and DNA binding proteins are rich in arginines and lysines, which comes as no surprise considering the high negative charge of RNA and DNA.[3] Histidine, lysine, and arginine residues also play important roles in the catalytic mechanisms of a large number of enzymes.[4−8] Thus, knowledge about the charge state and electrostatic interactions of basic amino acids is vital to the understanding of many interesting biological phenomena at the molecular level. Lysine and arginine have pK_a values of 10.4 and 12.0 in model compounds.[9] However, in proteins, lysine pK_a values have been observed to vary significantly. For example, in the calcium binding proteins calmodulin and calbindin D_{9k} they are found to vary[10,11] between 9.3 and 13.2. In special cases lysine pK_a values have been reported by indirect measurements to be downshifted to the extent that the amino groups titrate under physiologically relevant conditions. pK_a values as low as 6.0 and 9.1 have been determined. Using indirect techniques in the active sites of enzymes for lysine and arginine, respectively.[12,8] Methods that provide generally applicable and accurate *direct* measures of protonation constants of basic side chains with site-specific

† Lund University.
‡ University of Groningen.
§ Present address: Department of Biochemistry, box 357350, University of Washington, Seattle, WA 98195.

(1) Perl, D.; Mueller, U.; Heinemann, U.; Schmid, F. *Nat. Struct. Biol.* **2000**, *7*, 380−383.
(2) Strickler, S.; Gribenko, A.; Gribenko, A.; Keiffer, T.; Tomlinson, J.; Reihle, T.; Loladze, V.; Makhatadze, G. *Biochemistry* **2006**, *45*, 2761−2766.
(3) Fujita, M.; Kanehisa, M. *Genome Inform.* **2005**, *16*, 174−181.
(4) Borders, C. L., Jr.; Saunders, J. E.; Blech, D. M.; Fridovich, I. *Biochem. J.* **1985**, *230*, 771−776.
(5) Govindjee, R.; Misra, S.; Balashov, S. P.; Ebrey, T. G.; Crouch, R. K.; Menick, D. R. *Biophys. J.* **1996**, *71*, 1011−1023.
(6) Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M.; Escalante-Semerena, J. C. *Biochemistry* **2003**, *42*, 2866−2873.
(7) Sekimoto, T.; Matsuyama, T.; Fukui, T.; Tanizawa, K. *J. Biol. Chem.* **1993**, *268*, 27039−27045.
(8) Wells, T. N.; Scully, P.; Magnenat, E. *Biochemistry* **1994**, *33*, 5777−5782.
(9) Nozaki, Y.; Tanford, C. *Methods Enzymol.* **1967**, *11*, 715−734.
(10) Zhang, M.; Vogel, H. J. *J. Biol. Chem.* **1993**, *268*, 22420−22428.
(11) Kesvatera, T.; Jönsson, B.; Thulin, E.; Linse, S. *J. Mol. Biol.* **1996**, *259*, 828−839.
(12) Highbarger, L. A.; Gerlt, J. A.; Kenyon, G. L. *Biochemistry* **1996**, *35*, 41−46.

resolution would clearly be of value for mechanistic studies of enzyme active sites.

At the same time, pK_a values provide the most sensitive experimental probe of electrostatic interactions, and determination of pK_a values is an important means of studying electrostatic effects in proteins. Also, experimental pK_a values serve as a valuable basis for calibration of electrostatic calculations.[11,13,14] Calculations of pK_a values of lysine residues have previously been attempted and compared to experiments for calbindin D_{9k}[11] and protein GB1,[15] but method development has generally been hampered by the limited experimental data to benchmark against.

NMR spectroscopy is the most suitable and commonly employed method to obtain site-specific information on titrating groups.[16] Chemical shifts are exquisitely sensitive to changes in the electron distribution and electric fields, such that these can be monitored as a function of pH to provide information about the charge states and electrostatic coupling. Because of their prevalence, isotopic abundance, and inherent NMR sensitivity ^{1}H nuclei have traditionally been the reporter groups of choice in protein applications. Unfortunately, protons are rather weakly affected when they are located multiple bonds away from the titrating site and may respond to various charges developing in their environment. In a typical protein, with many charged groups, the response of ^{1}H chemical shift on pH is highly convoluted.[17,18] More reliable methods are therefore directed at monitoring ^{13}C or ^{15}N chemical shifts of the titrating groups in the side chains directly.[19,20] This method is feasible when specific or uniform labeling is employed to boost the natural isotope level, as is commonly done in modern biomolecular NMR spectroscopy. The inherently lower NMR sensitivity of ^{13}C or ^{15}N is overcome by indirect excitation and detection of attached or neighboring ^{1}H nuclei. For example, the pK_a values of carboxylate groups can be determined by monitoring side-chain carboxyl ^{13}C resonances in 2D H(C)CO experiments.[21,22] For histidines, the side-chain ^{15}N chemical shifts are correlated with the nonexchanging proton spins in the ring via two-bond ^{15}N$-^1$H correlation spectroscopy.[23]

Since 2D ^{15}N$-^1$H HSQC, HMQC, and TROSY experiments are among the most sensitive means to determine nitrogen-15 chemical shifts of backbone and side-chain amide groups in proteins these could be suitable for determining the pK_a of lysine and arginine as well. Indeed, at low pH and temperature arginine $^{15}N^{\epsilon}-^1H^{\epsilon}$ and lysine $^{15}N^{\zeta}-^1H^{\zeta}$ correlations may become detectable,[24,25] whereas arginine $^{15}N^{\eta}-^1H^{\eta}$ cross-peaks remain broadened by slow rotation of the guanidino head group.[26,27] Unfortunately, however, at elevated pH and temperature rapid hydrogen exchange precludes detection of amino protons, particularly for side chains that are exposed to solvent. For example, at 30 °C and pH 6.8 Farmer and Venters[28] observed none of the 24 expected cross-peaks for lysine residues in the HSQC spectrum of human carbonic anhydrase II. In another case Poon et al.[25] observed a single cross-peak at 30 °C and pH 6.5 for a buried lysine in the HSQC spectrum of *apo* CexCD out of a possible 18. About eight additional correlations could be observed at 10 °C with two more residues becoming observable when lowering the pH further to 5.8. No cross-peaks were observed when the pH was raised above 9. In a third application Iwahara et al.[29] were able to directly detect three lysine NH_3 cross-peaks of the HoxD9 homeodomain upon complexation to DNA at 35 °C, pH 5.8. Collectively these examples show that lysine amino groups are observable at (1) low pH and low temperature or (2) low pH in the context of hydrogen bonding or ion pairs. Due to the rapid exchange of amino protons with solvent water at high pH one-bond ^{15}N$-^1$H correlation spectroscopy is unsuitable as a general method for determination of lysine pK_a's. The high sensitivity of ^{1}H-based NMR methods that target the nonlabile protons in the side chain should be preferred. For chemically native lysine side chains a first general method for the site-specific determination of pK_a values has recently been reported based on incorporation of chemically synthesized [5-^{13}C]lysine into the bacterial growth medium during protein production.[30] The very high pK_a values of arginines in proteins cannot be determined due to protein instability and hydrolysis at high pH and inaccuracy of pH measurements above pH $\approx$ 12, unless severely downshifted.

Here, we present NMR techniques to determine the charge state of lysine (Lys) and arginine (Arg) residues that can be applied to uniformly ^{13}C/^{15}N-labeled proteins: titration events are monitored by the chemical shift of the side-chain nitrogen nuclei N^{ϵ} (Arg) and N^{ζ} (Lys). An experiment is described that yields the sequence-specific assignment of $^{15}N^{\epsilon}$ and $^{15}N^{\zeta}$ through correlations with the ^{13}C nuclei in the side chain, which are easily matched with a spectrum from a companion experiment starting from the backbone. For arginine the C^{ζ} shifts were monitored also. The N^{ϵ} and C^{ζ} nuclei are both expected to be reliable reporters of the guanidinium group in arginine because the head group charge is delocalized over the entire $N^{\epsilon}-C^{\zeta}-(N^{\eta}H_2)_2$ moiety.[26] Collectively, these experiments enable determination of site-specific pK_a values for individual lysine and arginine residues in proteins. In addition, N^{ϵ} (Arg) and N^{ζ} (Lys) may be used as sensitive probes of the local electrostatic environment around the side chains.

These new techniques allowed us to monitor the charge state of lysines and arginines in three proteins with different charge characteristics: Ca^{2+}-saturated calbindin D_{9k} (pI = 4.4, net charge −2), FKBP12 (pI = 9.6, net charge +1), and *apo* calmodulin (pI = 3.9, net charge −24). For the latter protein

(13) Juffer, A. H.; Vogel, H. J. *Proteins* **2000**, *41*, 554−567.
(14) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415−436.
(15) Khare, D.; Alexander, P.; Antosiewicz, J.; Bryan, P.; Gilson, M.; Orban, J. *Biochemistry* **1997**, *36*, 3580−3589.
(16) Szakacs, Z.; Kraszni, M.; Noszal, B. *Anal. Bioanal. Chem.* **2004**, *378*, 1428−1448.
(17) Forman-Kay, J. D.; Clore, G. M.; Gronenborn, A. M. *Biochemistry* **1992**, *31*, 3442−3452.
(18) Kesvatera, T.; Jönsson, B.; Thulin, E.; Linse, S. *Proteins* **2001**, *45*, 129−135.
(19) Zhu, L.; Kemple, M. D.; Yuan, P.; Prendergast, F. G. *Biochemistry* **1995**, *34*, 13196−13202.
(20) Grissom, C. B.; Markley, J. L. *Biochemistry* **1989**, *28*, 2116−2124.
(21) Oda, Y.; Yamazaki, T.; Nagayama, K.; Kanaya, S.; Kuroda, Y.; Nakamura, H. *Biochemistry* **1994**, *33*, 5275−5284.
(22) Qin, J.; Clore, G. M.; Gronenborn, A. M. *Biochemistry* **1996**, *35*, 7−13.
(23) Pelton, J. G.; Torchia, D. A.; Meadow, N. D.; Roseman, S. *Protein Sci.* **1993**, *2*, 543−558.
(24) Rao, N. S.; Legault, P.; Muhandiram, D. R.; Greenblatt, J.; Battiste, J. L.; Williamson, J. R.; Kay, L. E. *J. Magn. Reson., Ser. B* **1996**, *113*, 272−276.
(25) Poon, D.; Schubert, M.; Au, J.; Okon, M.; Withers, S.; McIntosh, L. *J. Am. Chem. Soc.* **2006**, *128*, 15388−15389.
(26) Henry, G. D.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 59−66.
(27) Mulder, F. A. A.; Spronk, C.; Slijper, M.; Kaptein, R.; Boelens, R. *J. Biomol. NMR* **1996**, *8*, 223−228.
(28) Farmer, B. T., II; Venters, R. A. *J. Biomol. NMR* **1996**, *7*, 59−71.
(29) Iwahara, J.; Jung, Y.; Clore, G. *J. Am. Chem. Soc.* **2007**, *129*, 2971−2980.
(30) Gao, G.; Prasad, R.; Lodwig, S. N.; Unkefer, C. J.; Beard, W. A.; Wilson, S. H.; London, R. E. *J. Am. Chem. Soc.* **2006**, *128*, 8104−8105.

the individual pK_a values of all lysine residues were determined and compared with previous studies based on chemical modification.

Materials and Methods

Expression and Purification. Uniformly labeled $^{13}C/^{15}N$ proteins were produced recombinantly in *Escherichia coli* using ^{15}N ammonium chloride and ^{13}C glucose as the sole nitrogen and carbon sources. Vertebrate calmodulin was purified as previously described.[31] *Apo* calmodulin was prepared by chelating Ca^{2+} with EDTA, and its subsequent removal on a gel filtration column. The production and purification of uniformly labeled $^{13}C/^{15}N$ calbindin D_{9k} P43G have been described elsewhere.[32] Recombinant human FKBP12 was purified on a CM cellulose column equilibrated with 10 mM Tris buffer, 1 mM EDTA, and 1 mM DTT by elution with a linear gradient of 0−0.5 M NaCl. Following lyophilization, the protein was desalted on a Sephadex G25 gel filtration column.

NMR Sample Preparation. All samples were dissolved in 90%/10% v/v $^1H_2O/^2H_2O$ with trace amounts of 2,2,-dimethylsilapentane-5-sulfonic acid (DSS) and sodium azide (NaN_3). *Apo* calmodulin samples were 2 mM protein in 100 mM potassium chloride unless indicated otherwise. The pH of the solution was adjusted by adding aliquots of concentrated NaOH in 100 mM KCl. The pH of the sample was measured before and after data collection, and an average value was used for data analysis. The NMR sample tube was filled with N_2 gas to avoid dissolution of carbon dioxide, which decreases the pH of the solution. The pH reading difference before and after acquisition was <0.1, except above pH 11, where larger changes where sometimes found. The pH was measured at room temperature using a Mettler Toledo MP225 pH meter connected to a Mettler Toledo U402-M3-S7/200 glass electrode. The pH was calibrated with pH buffers at 7.0, 9.21, 10.0, and 11.0 (Merck). Two buffers spanning the pH interval were used to calibrate the pH electrode. The sample of FKBP12 contained 1.5 mM protein dissolved in 25 mM potassium phosphate, pH 7.0. The sample of calbindin D_{9k} contained approximately 2 mM Ca^{2+}-loaded protein at pH 6.0.

NMR Experiments and Data Processing. All NMR spectra were acquired on 500 or 600 MHz Varian UNITY Inova spectrometers using triple-resonance probes with pulsed field-gradient capability. Data for calmodulin, FKBP12, and calbindin D_{9k} were collected at 27, 30, and 28 °C, respectively. In the H2(C)N and HD(CDNE)CZ experiments 60 and 512 complex points were collected in ω_1 (^{15}N or ^{13}C) and ω_2 (^1H), respectively. Spectral widths were 3000, 600, and 6000 Hz for ^{13}C, ^{15}N, and ^1H, respectively. For the titration experiments 16 or 32 transients were collected for each pH value. C-TOCSY-N(C)H2 spectra were acquired using 80, 40, and 512 complex points in ω_1, ω_2, and ω_3 covering spectral widths of 6000, 720, and 6000 Hz in the ^{13}C, ^{15}N, and ^1H dimensions, respectively. The C-TOCSY-(CO)NH experiment[33,34] was acquired using 80, 40, and 512 complex points in ω_1 (^{13}C), ω_2 (^{15}N), and ω_3 (^1H), covering spectral widths of 10 000, 1980, and 8000 Hz, respectively. ^1H chemical shifts were directly referenced to DSS, while ^{15}N and ^{13}C were indirectly referenced to ^1H using the conversion factors specified by the IUPAC standard.[35] Processing of NMR data was performed using the NMRPipe program suite,[36] and SPARKY (http://www.cgl.ucsf.edu/home/sparky) was used for assignment and peak analysis. All spectra were processed with a sine bell, typically shifted 90° in the indirect dimension and 60° in the acquisition domain. Zero filling increased the data size, and linear prediction was applied in the ^{15}N dimension of 3D experiments.

Data Analysis. Titration data were analyzed under the assumption that the titration sites are noninteracting and fitted using the Henderson−Hasselbalch equation adapted for NMR titrations

$$\delta_{obs} = \frac{(\delta_{HA} + \delta_{A-}\, 10^{(pH-pKa)})}{(1 + \delta_{A-}\, 10^{(pH-pKa)})} \quad (1)$$

where δ_{obs} is the observed chemical shift and δ_{HA} and δ_{A^-} are the chemical shifts of the protonated and unprotonated forms, respectively. No correction was made for isotope effects. Matlab (MathWorks Inc.) was used for nonlinear least-squares fitting of data. Errors in the pK_a value determination were determined by Monte Carlo sampling and represent fitting uncertainty alone.

H2CN (Lys/Arg) Correlation Experiments. 2D H2(C)N and H2C(N) correlation spectra were measured using the pulse sequence shown in Figure 1a. Narrow (wide) filled bars indicate 90° (180°) RF pulses applied with phase *x* unless otherwise indicated. Filled rectangular ^1H pulses were applied with a field strength of $\omega_1/2\pi = 35$ kHz. The ^1H carrier was centered at the water resonance (4.77 ppm) throughout. The open rectangle labeled "presat" represents weak ($\omega_1/2\pi = 20$ Hz) continuous wave (cw) irradiation to saturate the solvent water resonance during the relaxation delay of 1.0 s. Waltz-16 decoupling[37] was applied with $\omega_1/2\pi = 6.0$ kHz during nitrogen chemical shift evolution to suppress line broadening due to scalar relaxation resulting from proton exchange with the solvent.[38] Rectangular ^{15}N pulses with $\omega_1/2\pi = 6.0$ kHz were centered at 33 ppm, which is optimal for Lys while Arg has reduced intensity, or 59 ppm to detect Lys and Arg with similar sensitivity. The ^{15}N spectral width was set to ∼24 ppm so that peaks arising from Lys, Arg, and Asn/Gln side chains appear in different portions of the 2D map despite potential overlap in proton resonance frequencies. Backbone signals were suppressed using selective carbon pulses (see below), although peaks due to Gly were still visible. These signals were not found to be problematic since the proton shifts occurred downfield from the resonances of interest. Rectangular ^{13}C pulses were centered at 42 ppm and applied with $\omega_1/2\pi = 12.5$ (15.0) kHz at 500 (600) MHz up to point a and between b and c in the sequence. Between points a and b and between c and d rectangular ^{13}C 90° and 180° pulses were applied with $\omega_1/2\pi = 3.9$ (4.7) and 9.1 (10.9) kHz at 500 (600) MHz, respectively. The composite $90_y180_x90_y$ pulse between points b and c was applied with the carrier at 110 ppm and inverts symmetrically with side bands centered at 42 (Lys $^{13}C^{\epsilon}$, Arg $^{13}C^{\delta}$, Gln $^{13}C^{\gamma}$, Asn $^{13}C^{\beta}$) and 178 ppm (Arg $^{13}C^{\zeta}$, Gln $^{13}C^{\delta}$, Asn $^{13}C^{\gamma}$) given the power levels listed above. This decouples the side-chain nitrogen coherences from scalar coupled carbon nuclei. The sombrero shape indicates a band-selective REBURP 180° pulse[39] during the reverse INEPT transfer period. This pulse was applied with a peak field strength of $\omega_1/2\pi =$ 3.9 (4.7) kHz and pulse length of 1670 (1516) μs at 500 (600) MHz, respectively. At these power levels the REBURP pulse selectively refocuses (>98%) over a 20 ppm bandwidth. The carrier position of the pulse was shifted to 38 ppm by phase modulation so as to invert only the region upfield from 48 ppm. This region encompasses the side-chain Lys $^{13}C^{\epsilon}$, Arg $^{13}C^{\delta}$, Gln $^{13}C^{\gamma}$, and Asn $^{13}C^{\beta}$ chemical shifts but not those of backbone $^{13}C^{\alpha}$ other than Gly. Carbonyl inversion pulses were delivered at 180 ppm by phase modulation and serve to decouple Gln $^{13}C^{\gamma}$ and Asn $^{13}C^{\beta}$. These pulses can be omitted if Gln and Asn chemical shifts are not being monitored. ^{13}C decoupling during acquisition was achieved by GARP-1 decoupling[40] with $\omega_1/2\pi = 1.25$

(31) Waltersson, Y.; Linse, S.; Brodin, P.; Grundström, T. *Biochemistry* **1993**, *32*, 7866−7871.
(32) Mulder, F. A. A.; Akke, M. *Magn. Reson. Chem.* **2003**, *41*, 853−865.
(33) Logan, T. M.; Olejniczak, E. T.; Xu, R. X.; Fesik, S. W. *J. Biomol. NMR* **1993**, *3*, 225−231.
(34) Montelione, G. T.; Lyons, B. A.; Emerson, S. D.; Tashiro, M. *J. Am. Chem. Soc.* **1992**, *114*, 10974−10975.
(35) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wuthrich, K. *J. Biomol. NMR* **1998**, *12*, 1−23.
(36) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(37) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335−338.
(38) Abragam, A. In *Principles of Nuclear Magnetism*; Clarendon Press: Oxford, 1961.
(39) Geen, H.; Freeman, R. *J. Magn. Reson.* **1991**, *93*, 93−141.
(40) Shaka, A. J.; Barker, P. B.; Freeman, R. *J. Magn. Reson.* **1985**, *64*, 547−552.

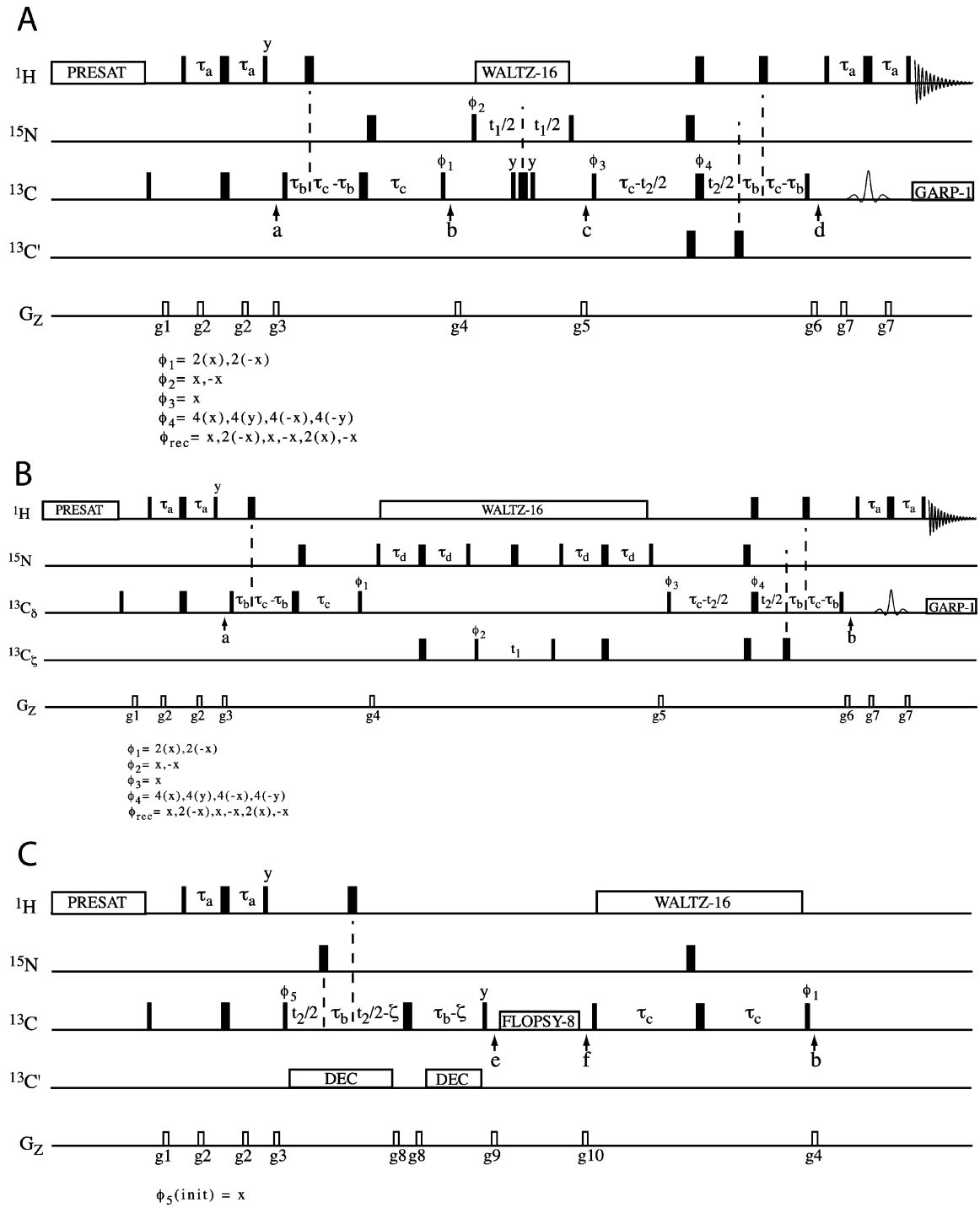


Figure 1. Pulse sequence for recording (a) H2CN spectra for arginine and lysine, (b) HD(CDNE)CZ spectra for arginine, and (c) C-TOCSY-N(C)H2 spectra of arginine and lysine.

kHz. The delays were as follows: $\tau_a = 1.7$ ms; $t_b = 1.9$ ms; and $\tau_c = 14.3$ ms. Gradient strengths in G/cm (lengths in ms) were as follows: $g_1 = 8.0$ (0.5); $g_2 = 5.0$ (0.5); $g_3 = 15.0$ (1.0); $g_4 = 10.0$ (0.2); $g_5 = -13.0$ (0.5); $g_6 = 16.0$ (0.5); and $g_7 = 8.0$ (0.5). Phase cycling: $\phi_1 = 2(x),2(-x)$, $\phi_2 = (x,-x)$, $\phi_3 = x,-x$ $\phi_4 = 4(x),4(y),4(-x),4(-y)$, $\phi_{rec} = (x),2(-x),(x),(-x),2(x),(-x)$. Phase-sensitive 2D H2(C)N and H2C(N) spectra were recorded using States-TPPI incrementation[41] of ϕ_2 and ϕ_3, respectively. Data sets were typically acquired with 128 complex data points in the indirect dimension and a spectral width of 24 ppm. The duration of each 2D experiment was 80 min.

The 2D H2(C)N experiment is also suitable for simultaneous detection of the N-terminal α-amino group by increasing the refocusing bandwidth to encompass the $^{13}C^{\alpha}$ shift of the first residue (~55 ppm). The protonation state of the α-amino group is then monitored from the $^{15}N^{\alpha}$ chemical shift. By increasing the refocusing bandwidth backbone correlations also become visible, although these are weak as a result of the large offset of the ^{15}N carrier (~33 ppm) from the backbone amide resonances (~120 ppm). However, because of the large shift separation, overlap between these signals

(41) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393−399.

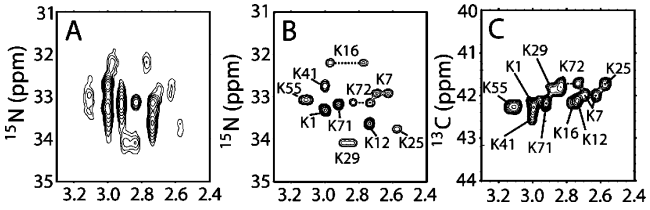


Figure 2. 2D H2CN spectra of calbindin D$_{9k}$. (a) H2(C)N spectrum with a central ^{1}H π pulse during t_1, (b) H2(C)N spectrum with composite pulse decoupling during t_1, (c) H2C(N) spectrum.

does not occur or can be avoided by a suitable choice of the spectral width.

HD(CDNE)CZ (Arg) Correlation Experiments. The 2D HD(CDNE)CZ correlation spectrum for *apo* calmodulin was measured using the pulse scheme shown in Figure 1b. The parameters were identical to those of the H2CN pulse sequence (Figure 1a) except for the following modifications. The ^{15}N carrier was placed at 80 ppm, between the Arg N$^\epsilon$ and N$^\eta$ resonances. The carbon carrier frequency was switched to ^{13}C$^\zeta$ (158 ppm) after g_4 and returned to 42 ppm after g_5, and the REBURP pulse was applied on-resonance. The additional transfer delay was $\tau_d = 12.5$ ms. The duration of the 2D experiment was 90 min.

C-TOCSY-N(C)H2 (Lys/Arg) Correlation Experiments. The 3D C-TOCSY-N(C)H2 correlation spectrum of FKBP12 was acquired at 600 MHz using the pulse scheme shown in Figure 1c. The final portion of this experiment is not displayed and is identical to the part of Figure 1a following g_4. The parameters were identical to those of the H2CN pulse sequence (Figure 1a) except for the following modifications. Rectangular ^{13}C pulses were centered at 42 ppm and applied with $\omega_1/2\pi = 15.0$ kHz up to point e. Between points f and b and between c and d rectangular ^{13}C 90° and 180° pulses were applied with $\omega_1/2\pi =$ 4.7 and 10.9 kHz. Isotropic ^{13}C mixing between points e and f was achieved through a FLOPSY-8 scheme[42] applied for 20 ms with $\omega_1/2\pi = 6.9$ kHz. Carbonyl ^{13}C decoupling can be achieved by SEDUCE decoupling[43] applied at 180 ppm through phase modulation with $\omega_1/2\pi = 0.9$ kHz. This was not implemented in the present experiment since the total evolution time was kept short enough (6 ms) that the effects of scalar coupling were insignificant. The delay ζ was increased from 0 to τ_b in concert with incrementing t_1 from 0 to $t_{1,max}$. This shared time evolution makes optimal use of the $2\tau_b$ period to encode chemical shift and avoids relaxation losses.[33] ^{15}N pulses were centered at 59 ppm. Additional gradients had the following strengths in G/cm (lengths in ms): $g_8 = 7.0$ (0.1); $g_9 = 10.0$ (0.4); and $g_{10} = 10.0$ (0.4). Phase-sensitive spectra were recorded using States-TPPI incrementation[41] of ϕ_5 and ϕ_3. The 3D data set covered spectral widths of 6000 × 720 × 6000 Hz using 80 × 40 × 512 complex points in the t_1 (^{13}C), t_2 (^{15}N), and t_3 (^{1}H) dimensions, respectively. The total experiment time was 70 h.

Assignment of *apo* Calmodulin at High pH. Assignment of the ^{15}N$^\zeta$ resonances in the H(C)N spectrum of *apo* calmodulin was attempted but proved to be very difficult. At neutral pH, six of the eight lysine residues have identical shifts for all of the side-chain ^{1}H$^\epsilon$, ^{13}C$^\epsilon$, and ^{15}N$^\zeta$ resonances, which indicates that the chemical environment is highly similar for these residues. Only at higher pH values, where the different lysines start to titrate, is the chemical shift dispersion of ^{15}N$^\zeta$ sufficient to observe individual resonances (Figure 6). On the basis of the chemical shift dispersion in ^{15}N$^\zeta$ at higher pH values, the C-TOCSY-N(C)H2 experiment should provide the necessary correlations in principle. However, no clear cross-peaks were present in the 3D spectrum, while signal was observed in the individual 2D planes. This discrepancy is probably due to a drift in pH over the 3 day duration of the 3D experiment. At higher pH values dissolution of carbon dioxide leads to significant drift in pH. Even if stable sample conditions can be obtained it is questionable if the dispersion found at higher pH will be enough to assign all lysine residues in a C-TOCSY-N(C)H2 experiment. Due to exchange with solvent protons experiments based

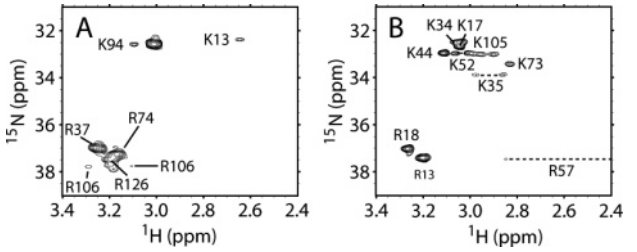


Figure 3. 2D H2(C)N spectra for (a) *apo* calmodulin and (b) FKBP12. Prochiral protons are connected by a dashed line. Signals of Arg N$^\epsilon$ appear aliased to lower chemical shift by 47.4 ppm.

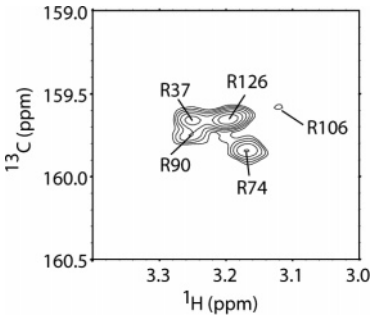


Figure 4. HD(CDNE)CZ spectrum of *apo* calmodulin.

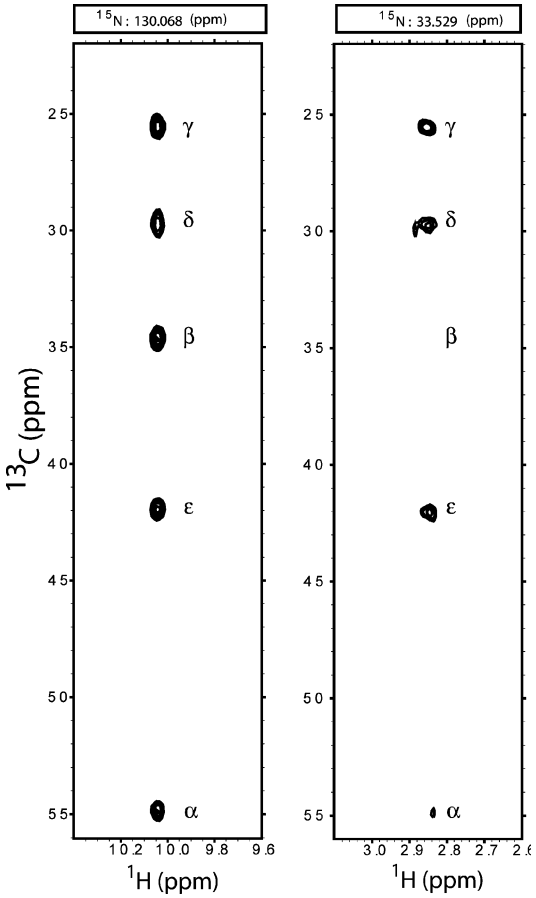


Figure 5. Assignment of lysine and arginine residues of FKBP12 from C-TOCSY-(CO)NH (left) and C-TOCSY-N(C)H2 (right).

(42) Mohebbi, A.; Shaka, A. J. *Chem. Phys. Lett.* **1991**, *178*, 374−378.
(43) McCoy, M. A.; Mueller, L. *J. Am. Chem. Soc.* **1992**, *114*, 2108−2112.

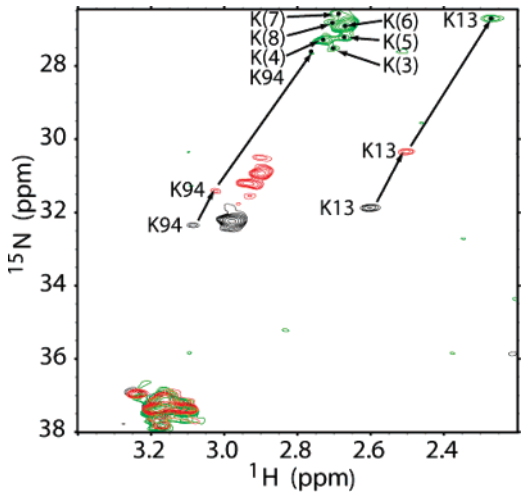


Figure 6. Titration of *apo* calmodulin. Superposition of spectra obtained at pH 9.67 (black), 10.44 (red), and 11.44 (green).

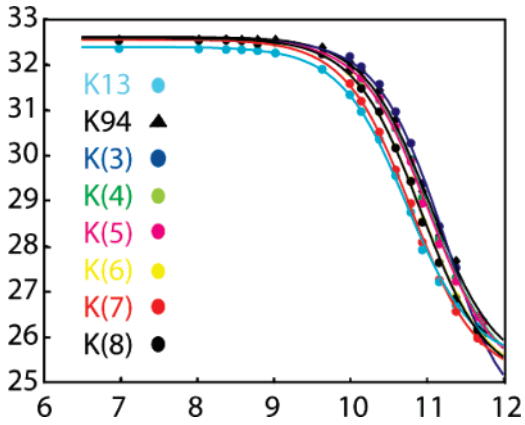


Figure 7. Titration of lysine residues in *apo* calmodulin followed by the change in N^{ξ} chemical shift. K(*X*), $X = 3-8$, represents residues that were not assigned sequence specifically.

on the backbone NH group will show few or no correlations. Also, the limited dispersion in H^{ϵ} and C^{ϵ} makes these chemical shifts useless in differentiating between different lysine residues in HCCH-TOCSY experiments. K13 and K94 were identified on the basis on their H^{ϵ} shifts and published assignment of *apo* calmodulin.[44] K13 (and to a lesser extent K94) is shifted in the ^{1}H dimension because of nearby ring currents. The family of NMR structures of *apo* calmodulin (1DMO) indicates that F65 is within 3−4 Å from the Hϵ protons. Although no clear close contact of K94 with an aromatic group is evident from 1DMO (F92 is closest but more than 6 Å away), K94 is the only other lysine residue in proximity of an aromatic group in the calcium-loaded structure (1PRW). Correlations for all lysine residues at neutral pH were present in C-TOCSY-(CO)NH with similar intensities, indicating that all lysines are present in the C-TOCSY-N(C)H2 experiment also. Six overlapped correlations can indeed be discerned in the H2(C)N spectra, which become resolved at pH between 10 and 12. Although the identity of individual residues could not be established as described above, the six correlations clearly belong to the remaining lysine side chains. Since the acidity constants for these residues were so similar (within 0.4 pK_a units) no further attempts were made to assign the H2(C)N spectrum at high pH, but the lysines were simply numbered 3−8 in Figure 7. All six arginines are identified in the C-TOCSY-(CO)NH spectra, but R86 is missing from the H2C(N), H2(C)N, and H2(CDNE)CZ spectra. H^{δ}, C^{δ}, and N^{ϵ} could be assigned by combining the C-TOCSY-(CO)NH, H2(CDNE)CZ, and H2(C)N spectra, see Figures 3a and 4.

(44) Zhang, M.; Tanaka, T.; Ikura, M. *Nat. Struct. Biol.* **1995**, *2*, 758−767.

Results

NMR Pulse Sequences To Follow Lys and Arg Protonation States. Three NMR pulse sequences were developed to directly monitor the titrating groups of Lys and Arg residues and assign the studied resonances. These are introduced briefly in this section and discussed more extensively in the following. The H2CN experiment is optimized for the study of Lys by correlating the nonexchanging H^{ϵ} proton with N^{ζ}, which is directly bound to the titrating proton. At the same time correlations between H^{δ} and N^{ϵ} for Arg are observed and the chemical shift of N^{ϵ} can be used to follow the titration of Arg. The experiment can be run as a 3D experiment if overlap is severe, but for the protein applications described below 2D H(C)N or HC(N) planes were found sufficiently resolved and time efficient. To follow the charge state of the Arg guanidinium group as a consequence of (de)protonation the 2D H2(C)N experiment was extended to correlate H^{δ} and C^{ζ} in the HD(CDNE)CZ experiment. Importantly, these experiments rely on detection at nonexchanging carbon-bound protons, so that these experiments can be performed over the full pH range and at ambient or elevated temperature without signal loss or broadening due to exchange with solvent water. In addition, to obtain residue-specific assignments of the H^{ϵ}/N^{ζ} (Lys) and H^{δ}/N^{ϵ} (Arg) chemical shift correlations in the side chains of Arg and Lys we developed an assignment experiment which correlates side-chain nitrogens with carbon and proton chemical shifts that is complementary to C-TOCSY-(CO)NH.[33,34]

Briefly, the H2CN experiment makes use of magnetization transfer from $^{1}H^{\delta} \rightarrow {}^{13}C^{\delta} \rightarrow {}^{15}N^{\epsilon}$ (Arg) and $^{1}H^{\epsilon} \rightarrow {}^{13}C^{\epsilon} \rightarrow {}^{15}N^{\zeta}$ (Lys) through one-bond scalar couplings, similar to the implementation of Yamazaki et al.,[45] with a number of additions and modifications. Most importantly, continuous pulse decoupling, rather than a central ^{1}H inversion pulse, accomplishes refocusing of the $^{1}H-{}^{15}N$ scalar coupling during the ^{15}N evolution period. This is crucial for application to lysine residues since excessive line broadening due to solvent exchange renders these spectra unintelligible otherwise: Chemical exchange on the time scale of several hundreds per second acts as destructive interference and line broadening, known as "scalar relaxation of the second kind".[38] Decoupling by a central refocusing pulse yielded the spectrum shown in Figure 2a, compared to Figure 2b when continuous decoupling was applied. Correlations in the backbone—which in principle are also observable in this experiment—are selected against by positioning the ^{13}C and ^{15}N carriers in spectral regions optimal for side-chain excitation and using selective refocusing. In the present case we have chosen to select for side-chain correlations by introducing a band-selective ^{13}C pulse in the final reverse INEPT step. Alternatively, the sequence could easily be adapted to deliver band-selective ^{15}N 180° pulses for Arg $^{15}N^{\epsilon}$ (~86 ppm) or Lys $^{15}N^{\zeta}$ (~33 ppm) only. (The specific implementation chosen here was motivated simply by the availability of only two waveform generators, dedicated to ^{1}H and ^{13}C in our setup.) Therefore, in the spectra correlations for Asn and Gln side chains are also observable due to the large

(45) Yamazaki, T.; Yoshida, M.; Nagayama, K. *Biochemistry* **1993**, *32*, 5656−5669.

two-bond scalar couplings between $^{13}C^{\gamma}-^{15}N^{\epsilon}$ in Gln and $^{13}C^{\beta}-^{15}N^{\delta}$ in Asn.

When performed in 2D mode, the pulse sequence of Figure 1a enables selective detection of $^1H-^{15}N$ or $^1H-^{13}C$ correlation spectra of nonaromatic nitrogen-containing side chains. Figure 2b and c shows for calbindin D_{9k} that even though basic amino acid side chains are generally on the protein surface, such spectra are quite well resolved. The main motivation to record $^1H-^{15}N$ rather than $^1H-^{13}C$ maps is that the response of the nitrogen nuclei to protonation of Lys are very large, about 7 ppm.[19] Fortuitously, the resolution of $^{15}N^{\zeta}$ chemical shifts of Lys is also superior to that of $^{13}C^{\epsilon}$ (compare Figure 2b and c). In addition, H2(C)N experiments are preferred over H2C(N) since potential spectral overlap of Arg H^{δ} and Lys H^{ϵ} is easily resolved along the ^{15}N dimension, whereas the chemical shift separation of $^{13}C^{\delta}$ in Arg and $^{13}C^{\epsilon}$ in Lys is much smaller. 2D $^1H-^{15}N$ spectra for *apo* calmodulin and FKBP12 (Figure 3) demonstrate that the resolution of the H2(C)N spectrum is expected to be sufficient for protein applications in general.

Detection of the N-terminal α-amino group was also done with the H2(C)N pulse sequence of Figure 1a, replacing the final shaped carbon inversion pulse with a full power rectangular 180° pulse. The resulting spectrum (Figure S5 in Supporting Information) is very similar to Figure 2b but contains additional cross-peaks. With the ^{15}N carrier placed at 33 ppm, signals due to the backbone were observed, but these were weak and aliased. The signal due to the α-amino group appeared at 38.0 ppm. Although possible, the protonation state of the N-terminus was not monitored in the current application.

The H2(C)N experiment was extended in a straightforward fashion with an additional transfer step to record Arg $^{13}C^{\zeta}$ chemical shifts. Since the charge of the guanidinium group is delocalized over the entire $N^{\epsilon}-C^{\zeta}-(N^{\eta}H_2)_2$ moiety,[26] the $^{15}N^{\epsilon}$ and $^{13}C^{\zeta}$ shifts should both be sensitive reporters of the side-chain protonation state. Figure 1b shows the pulse sequence for the HD(CDNE)CZ experiment in which magnetization transfer follows the pathway $^1H^{\delta} \rightarrow {}^{13}C^{\delta} \rightarrow {}^{15}N^{\epsilon} \rightarrow {}^{13}C^{\zeta}$, thereby correlating $^1H^{\delta}$ with $^{13}C^{\zeta}$. The spectrum for *apo* calmodulin, shown in Figure 4, demonstrates good sensitivity. All peaks for the six arginines in calmodulin are observed, but at least two signals overlap, as can be deduced from combining Figures 3a and 4.

Assignment of Lys N^{ζ} and Arg N^{ϵ} Chemical Shifts. In order to sequence-specifically assign the side-chain H^{ϵ}/N^{ζ} (Lys) and H^{δ}/N^{ϵ} (Arg) correlations, we developed a 3D experiment that correlates the side-chain nitrogen shifts with the protons on the adjacent carbon, as in H2(C)N, and with all carbon shifts of the side chain using isotropic mixing. For example, for Lys the magnetization transfer follows the pathway $H^{ali} \rightarrow C^{ali}$ (t_1) $\rightarrow C^{\epsilon} \rightarrow N^{\zeta}$ (t_2) $\rightarrow C^{\epsilon} \rightarrow H^{\epsilon}$ (t_3), where ali indicates all aliphatic nuclei of the side chain. Although experiments exist for the assignment of Arg N^{ϵ} resonances using the attached proton,[24,28] such experiments do not work when exchange with the solvent is rapid. These conditions prevail at room temperature for solvent-exposed Lys at most pH values, and for Arg at pH higher than about 6, signals are not observable. The C-TOCSY-N(C)-H2 pulse sequence (Figure 1c) was applied to FKBP12, and the results are shown in Figure 5. A C-TOCSY-(CO)NH spectrum was also recorded, and the residue-specific side-chain assignments were easily extended to the Lys $^{15}N^{\zeta}$ as shown.

Table 1. Chemical Shifts of Lysine $^{15}N^{\xi}$ for FKBP12 and Calbindin D_{9k}

FKBP12		calbindin D_{9k}	
residue	$^{15}N^{\xi}$ (ppm)	residue	$^{15}N^{\xi}$ (ppm)
K17	32.8	K1	32.1
K34	32.6	K7	32.2
K35	34.0	K12	33.0
K44	33.1	K16	31.5
K47		K25	33.1
K52	33.1	K29	33.4
K73	33.6	K41	32.6
K105	33.1	K55	32.4
		K71	32.5
		K72	32.5

Using this strategy the $^{15}N^{\zeta}$ chemical shifts were obtained for all seven lysines visible (out of eight expected from the sequence) in the H2(C)N spectrum for FKBP12. In the same manner complete Lys $^{15}N^{\zeta}$ assignments were obtained for calbindin D_{9k}. The shifts for the two proteins are compiled in Table 1.

Comparison of Spectra for *apo* Calmodulin, FKBP12, and Calbindin D_{9k}. The H2(C)N spectrum of *apo* calmodulin represents a most challenging scenario where six of the lysine residues are found at the exact same chemical shift for both H^{ϵ} and N^{ξ} (Figure 3a). In the case of Arg the situation is slightly better with five of six possible $H^{\delta}-N^{\epsilon}$ correlations detected. The spectrum of FKBP12 (Figure 2c) is significantly more dispersed, making it possible to identify seven out of eight lysine residues, all of which could easily be assigned using the C-TOCSY-N(C)H2 experiment, see Table 1. The situation for calbindin D_{9k} is also favorable, and here all 10 lysine residues were found and assigned (Figure 2b and Table 1). Taken together, 23 of 25 lysine residues were present in a data set comprising three different proteins, clearly demonstrating that H2(C)N spectra are suitable for monitoring the charge states of lysine residues.

Titration of *apo* Calmodulin. *apo* calmodulin was titrated in the pH range from 6.5 to 11.84, and 2D $^1H^{\epsilon}-^{15}N^{\zeta}$ correlations were recorded at each titration point. At neutral pH six out of eight lysine residues have an identical chemical shift in both the ^1H and ^{15}N dimensions (Figure 3a), indicating that they have a similar chemical environment. As the lysines start titrating at higher pH, the chemical shift dispersion increases so that individual resonances are visible in the spectra and the pK_a values for all lysines can be determined (although chemical shift assignments were verified for only two of them, as outlined in Materials and Methods). Titration curves for $^{15}N^{\zeta}$ are shown in Figure 7. Fitting of eq 1 to the experimental data yields total chemical shift differences between protonated and unprotonated forms of 7.5 ± 0.3 and 0.4 ± 0.06 ppm for $^{15}N^{\zeta}$ and $^1H^{\epsilon}$, respectively. This is consistent with the difference in chemical shift observed for the titration of lysines in mellitin observed by direct ^{15}N spectroscopy,[19] where shift differences of N^{ζ} between protonated and unprotonated forms were around 7.5 ppm.

The experimentally determined pK_a values of *apo* calmodulin are listed in Table 2. The pK_a values for *apo* calmodulin appear upshifted by 0.5 units, on average, compared to the values predicted from model compounds.[9] This is expected because of the high net negative charge (−24 at neutral pH) of the protein. The determined pK_a values fall into a narrow range

Table 2. pK_a Values of Lysine Residues in *apo* Calmodulin As Determined from N^{ζ} or H^{ϵ} at 25 °C in 100 mM KCl; K(*x*) Indicates That the Residue Is Not Assigned

residue	N^{ζ}	H^{ϵ}	$N^{\zeta}+H^{\epsilon}$
K13	10.74 ± 0.02	10.72 ± 0.06	10.74 ± 0.02
K94	11.03 ± 0.03	11.21 ± 0.08	11.05 ± 0.04
K(3)	11.22 ± 0.03	11.14 ± 0.14	11.22 ± 0.05
K(4)	11.04 ± 0.02	11.10 ± 0.09	11.04 ± 0.02
K(5)	11.03 ± 0.02	11.05 ± 0.07	11.02 ± 0.02
K(6)	10.92 ± 0.01	10.81 ± 0.07	10.93 ± 0.02
K(7)	10.80 ± 0.01	10.81 ± 0.08	10.80 ± 0.02
K(8)	10.93 ± 0.02	10.95 ± 0.08	10.93 ± 0.03

with the largest difference in pK_a being 0.5 units. All titration curves display ideal two-state equilibrium behavior, and good fits are obtained without introducing a Hill parameter, which is sometimes necessary to account for "nonideal" titration curves.[46] Due to low signal-to-noise at pH above 12 from the repeated addition of base solution, the titration did not yield a good definition of the chemical shift for the neutral species. Nevertheless, the pK_a values could be determined very precisely (within 0.1). Note that the uncertainties listed in Table 2 only represent the fitting error, while the largest error is expected to arise from uncertainty of the pH at larger values. This accounts for an error on the order of 0.1. Notwithstanding the absolute error in the pK_a values, differences in acidity constants obtained from monitoring the large change in $^{15}N^{\zeta}$ chemical shift may be as small as 0.01, allowing extremely small effects of electrostatic environments to be quantified.

Calmodulin has previously been shown to be stable and maintain its structure at high pH values, even in the absence of bound calcium.[47] We preformed additional experiments to confirm that the protein maintained its native structure and composition at high pH values. Circular dichroism (S1), $^{15}N-^{1}H$ and $^{13}C-^{1}H$ HSQC (S2 and S3), and agarose gel electrophoresis (S4) were used to demonstrate that calmodulin is stable, maintains its structure, and resists deamidation at high pH values on the time scale of the titration experiments.

Discussion

pK_a values are the most accurate experimental probe of electrostatic interactions in proteins, and NMR is the most generally applicable technique to obtain site-specific information on the titration status of a charged residue. For Lys, Asp, and Glu, 2D-[^{1}H,^{1}H]-TOCSY, 2D-[^{15}N,^{1}H]-HSQC-TOCSY, or 2D-[^{1}H,^{1}H]-NOESY experiments have been used extensively to determine pK_a values.[15,18,48−50] Although straightforward in principle, these techniques suffer from a number of practical drawbacks. Signal overlap is often a considerable problem, especially for larger proteins, due to limited chemical shift dispersion in the ^{1}H dimension. More seriously, ^{1}H chemical shifts may sense multiple protonation equilibria simultaneously, resulting in complex titration curves that may be difficult to interpret. Because of their sensitivity to structure, protons may sometimes even report on rearrangements rather than an actual titration event. It is not uncommon that protons attached to the same carbon atom report on different pK_a values due to the influence of nearby titrating groups or other changes in the microenvironment.[17,48,49,51] These problems can be avoided to a large degree by following the chemical shift of heteronuclei at or near the titrating group. For example, the ^{13}C carboxyl shifts accurately report on the electrostatics of aspartic and glutamic acids.[21] The determined pK_a value is a reliable estimate of the actual pK_a value because of the changes in electron density and resonance structure, transmitted through chemical bonds. In the case of lysines, a number of methods have been used to determine pK_a values. These include indirect methods where functional aspects of the protein are followed. For example, pK_a values have been obtained by monitoring the rate of catalysis or inactivation of an inhibitor as a function of pH.[8,12] These types of methods do not provide site-specific information and will be useful only in cases where a limited number of titrating residues contribute to the observed effect and only for proteins that have a pH-dependent property amenable to characterization. The pH-dependent stability of a protein can be interpreted in terms of pK_a shifts between the folded and unfolded state so that the average charge state of the lysine residues can be determined. Selective ^{13}C labeling of lysine residues has been used to obtain pK_a values by following 2D-[H^{ϵ},C^{ϵ}][11] or 2D-[H^{δ},C^{δ}][30] correlations as a function of pH. The drawback of this method is that the carbon chemical shift changes are small and potentially influenced by titrating groups in the surrounding area. Moreover, the selectively labeled protein sample is of limited use for other applications. Selective labeling with ^{15}N has been used in determination of pK_a values in mellitin by ^{15}N spectroscopy.[19] However, assignment of the ^{15}N resonances required site-specific labeling through solid-phase synthesis[19] or that the lysines are removed selectively by mutation. In another method, the lysine residues are chemically modified with ^{13}C-labeled mono- or dimethyl groups, as demonstrated for calmodulin[10] and calbindin D_{9k}.[52] This method yields simple spectra, but there is great risk that the chemical modification perturbs the time-averaged protein structure and induces bias in the population of side-chain conformers, both of which change the pK_a values of the modified residues. In addition, methylation induces an offset in the pK_a. For these samples the assignment problem remains the same as for selectively ^{15}N-labeled lysine residues mentioned above.

To supersede existing methodology we developed an NMR method to determine pK_a values of lysine and arginine residues, which can be applied to uniformly $^{15}N/^{13}C$-labeled proteins. In this method 2D-[H^{ϵ},N^{ζ}] (Lys) and 2D-[H^{δ},N^{ϵ}] (Arg) correlations are recorded in a single experiment as a function of pH. Since nonexchangeable protons are used for excitation and detection these experiments are applicable over the full range of pH and temperature. In the case of arginine the head group charge is delocalized over the entire $N^{\epsilon}-C^{\zeta}-(NH_2)_2$ moiety, such that N^{ϵ} is a suitable reporter for the arginine charge state. For lysine the N^{ζ} nucleus is directly bonded to the titrating proton, which yields significant chemical shift changes between the protonated and unprotonated forms. As chemical shifts can be measured with great precision this results in very precise data for the titration event. Furthermore, high accuracy is achieved because

(46) Markley, J. L. *Acc. Chem. Res.* **1975**, *8*, 70−80.
(47) Huque, M. E.; Vogel, H. J. *J. Protein Chem.* **1993**, *12*, 695−707.
(48) Perez-Canadillas, J. M.; Campos-Olivas, R.; Lacadena, J.; Martinez, del Pozo, A.; Gavilanes, J. G.; Santoro, J.; Rico, M.; Bruix, M. *Biochemistry* **1998**, *37*, 15865−15876.
(49) Song, J.; Laskowski, M., Jr.; Qasim, M. A.; Markley, J. L. *Biochemistry* **2003**, *42*, 6380−6391.
(50) Sundd, M.; Iverson, N.; Ibarra-Molero, B.; Sanchez-Ruiz, J. M.; Robertson, A. D. *Biochemistry* **2002**, *41*, 7586−7596.
(51) Lindman, S.; Linse, S.; Mulder, F. A.; Andre, I. *Biochemistry* **2006**, *45*, 13993−14002.
(52) Zhang, M.; Thulin, E.; Vogel, H. J. *J. Protein Chem.* **1994**, *13*, 527−535.

the change in chemical shift will almost exclusively report on the titration of the residue itself and not on secondary events, such as titration of other groups in the proximity or changes in the local magnetic environment. This is in contrast to titration data based on ^{1}H resonances, which are often convoluted.[17,49] The dispersion in $^1H^{\epsilon}-^{15}N^{\zeta}$ makes it possible to monitor selectively lysine residues in moderate size proteins where severe overlap occurs for ^{1}H resonances.

The situation in *apo* calmodulin represents a worst-case scenario, where six out of eight $^1H^{\epsilon}-^{15}N^{\zeta}$ cross-peaks have virtually the same chemical shifts in both dimensions (Figures 3a and 6). Even in this case, the titration can be monitored selectively for all lysine residues since the chemical shift in ^{15}N is significantly different at higher pH values. The dispersion is significantly better in FKBP12 and calbindin D_{9k} (Figures 2 and 3c). In these two proteins all lysine and arginine $^1H^{\epsilon}-^{15}N$ resonances can be resolved, and a number of prochiral $^1H^{\epsilon}$ also have different chemical shifts, indicating that they are not fully solvent exposed and/or are dynamically averaged. In total, we observe 23 out 25 lysine residues in the $^1H^{\epsilon}-^{15}N^{\zeta}$ correlation maps, which clearly demonstrate the feasibility of the method to study the charge state of basic residues. We further developed a method to assign $^{15}N^{\zeta}$ resonances in lysines and $^{15}N^{\epsilon}$ resonances in arginines, as demonstrated for FKBP12 (Figure 5 and Table 1).

pH Titration of *apo* Calmodulin. We applied the H2(C)N experiment to determine lysine pK_a values in *apo* calmodulin, which maintains its structure at high pH (Supporting Information, Figures S1−S4). The pK_a values of *apo* calmodulin are, on average, upshifted by 0.5 units compared to model values, as expected from the large net negative protein charge. The magnitude of the shift is however limited by electrostatic screening of 100 mM KCl. For reference, the pK_a values of lysines in calbindin D_{9k} are reduced by 0.66 units in 100 mM salt which can be seen by combining data by Kesvatera[11] and Zhang.[52] The chemical shift correlations of $^1H^{\delta}-^{15}N^{\epsilon}$ of arginine side chains are monitored at the same time as $^1H^{\epsilon}-^{15}N^{\zeta}$ of lysine. We observe small chemical shift changes for arginine at the highest pH values, indicating that the pK_a values lie above 12.5 in *apo* calmodulin. This is expected since lysine residues also demonstrated upshifted pK_a values, and the arginines in calmodulin will titrate in an even more negative environment than the lysines.

In summary, for calmodulin electrostatic coupling between lysine residues appears to be very modest in the absence of bound calcium. This conclusion is based on the following observations: First, the small pK_a shifts relative to model values indicate that all lysine side chains are fully solvent exposed and interact through high dielectric environments. Second, the uniform chemical shifts of the lysine residues indicated a lack of specific interactions. Third, the largest difference between individual lysine pK_a values is 0.5 units, showing that all monitored lysine residues experience highly similar electrostatic environments. Fourth, the data could be fitted perfectly to the Henderson−Hasselbalch model without the necessity for a Hill parameter, which shows that the lysines residues have weak electrostatic coupling and indicates that the diversity in pK_a values is primarily determined by differences in interactions with carboxyl groups.[51]

The pK_a values obtained in this study can be compared with those determined using ^{13}C-labeled methylated lysines.[10] The values for the dimethylated lysines in the study of Zhang and Vogel[10] are systematically lower by 0.3−0.9. This is in agreement with other studies of reductively alkylated proteins and peptides.[53] For example, for free lysine addition of two methyl groups lowers the pK_a by 0.6 units compared to the unmodified group. Of note, there is more spread in the values obtained for the modified lysine residues by ^{13}C NMR than observed here for the unreacted protein using ^{15}N NMR. In particular, the protonation constant for dimethylated Lys75 (9.87) was found to be significantly lower than those for the other lysines (10.23−10.55), a behavior not found in our study of recombinant calmodulin, where none of the lysine groups is modified. Whether and how these differences arise from chemical modification awaits further investigation.

Conclusions

Electrostatic interactions in proteins can be probed experimentally through determination of residue-specific acidity constants. As the spin nuclear shielding is exquisitely sensitive to the surrounding electron distribution, pK_a values can be assigned to individual chemical groups. Because of their prevalence in proteins and for sensitivity ^{1}H spins have been most commonly employed in NMR spectroscopic investigations of protonation equilibria. Unfortunately, the monitored ^{1}H nuclei are generally separate from the titration site and, as a consequence, their chemical shifts often report on multiple nearby titration events. We have shown here that truly individual pK_a values can be obtained for lysine and arginine by following heteronuclear—^{13}C and ^{15}N—chemical shifts in protein side chains. The experiments proposed have excellent sensitivity since they employ ^{1}H excitation and detection and should be generally applicable to doubly labeled proteins commonly produced for structure determination. The present set of experiments is not limited to their use to determination of pK_a values of lysine and arginine residues. The reporter groups used in this study are sensitive to the chemical environment around the side chain and may prove useful for other processes, such as ligand binding, protein−protein interactions, or enzyme catalysis.

Acknowledgment. We thank Ulrika Brath for production and purification of FKBP12 and Eva Thulin for purification of *apo* calmodulin. Mikael Akke is acknowledged for his continuous support.

Supporting Information Available: Data showing that *apo* calmodulin retains its structure and is not deamidated at high pH; H2(C)N spectrum of calbindin D_{9k}, recorded to include the N-terminal backbone amino group. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0721824

(53) Means, G. *J. Protein Chem.* **1984**, *3*, 121−130.