Quantitative Identification of the Protonation State of Histidines in Vitro and in Vivo†

Nobuhisa Shimba, Zach Serber, Richard Ledwidge, Susan M. Miller, Charles S. Craik, and Volker Dötsch*

Departments of Pharmaceutical Chemistry and Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, California 94143

Received March 24, 2003; Revised Manuscript Received April 22, 2003

ABSTRACT: The C−N coupling constants centered at the $C^{\epsilon 1}$ and $C^{\delta 2}$ carbons in histidine residues depend on the protonation state and tautomeric form of the imidazole ring, making them excellent indicators of pH or pK_a, and the ratio of the tautomeric states. In this paper, we demonstrate that the intensity ratios for the $C^{\epsilon 1}$-H and $C^{\delta 2}$-H cross-peaks measured with a constant time HSQC experiment without and with J_{C-N} amplitude modulation are determined by the ratios of the protonated and deprotonated forms and tautomeric states. This allows one to investigate the tautomeric state of histidines as well as their pK_a in situations where changing the pH value by titration is difficult, for example, for in-cell NMR experiments. We apply this technique to the investigation of the bacterial protein NmerA and determine that the intracellular pH in the *Escherichia coli* cytoplasm is 7.1 ± 0.1.

Histidine is frequently involved in the function of proteins, mainly because of the chemical versatility of its imidazole ring, which includes protonated and deprotonated forms as well as the tautomeric states (*1*−*3*). A recent analysis of the enzyme structural database (www.biochem.ucl.ac.uk/bsm/enzymes) revealed that approximately 50% of all enzymes use histidine in their active sites. To determine the protonated and deprotonated forms and the tautomeric states of histidine residues in solution, NMR spectroscopy has been extensively used with good success. Traditionally, the pK_a of histidines has been determined by measuring chemical shift changes during pH titration experiments. However, this method cannot be applied to proteins that are unstable over a wide pH range, and it does not provide information about the tautomeric states. To circumvent this problem, several NMR techniques have been developed, based on HMQC[1] (*4*−*6*), HMBC[1] (*7*, *8*), and HSMQC[1] (*9*, *10*) experiments, which discriminate each state from the ratio of two- and three-bond remote couplings, such as $^2J_{N\epsilon 2-H\delta 2}$ and $^3J_{N\delta 1-H\delta 2}$ in the imidazole ring. Unfortunately, the small value of these couplings necessitates long coherence transfer times, making this method not applicable to large proteins. The method can, however, be extended to proteins of higher molecular weight if larger one-bond couplings centered at the $C^{\epsilon 1}$ and $C^{\delta 2}$ carbons in the imidazole ring are used (*11*, *12*). It was shown previously that these large coupling constants can be used to characterize the protonation and tautomeric state of all seven histidine residues in a Fab fragment of the catalytic antibody 6D9 (MW ∼50 000) (*11*). However, so far the C−N coupling constants of the histidine residues have only been used for a qualitative analysis to identify the state with the highest population. In this paper, we show that these coupling constants can also be used to provide quantitative analysis of the different histidine states. In addition, we show that by measuring the ratio of each state, we can also determine the pH in living cells by in-cell NMR experiments (*13*−*17*) since the typical pK_a values of histidine residues range from 5.0 to 8.0, which corresponds to the range of pH values so far shown to exist in cells.

EXPERIMENTAL PROCEDURES

Materials. Uniformly $^{13}C/^{15}N$ labeled histidine and $^{13}C/^{15}N$ labeled media were purchased from Cambridge Isotope Laboratories, Inc. Uniformly $^{13}C/^{15}N$ labeled histidine was dissolved in 50 mM sodium formate at pH 3.3, 3.7, and 4.2; 50 mM sodium acetate at pH 5.2; 50 mM MES at pH 5.7 and 6.2; 50 mM sodium phosphate at pH 6.7, 7.2, and 7.7; and 50 mM BICINE at pH 8.2 and 8.7, respectively. The final concentration of each uniformly $^{13}C/^{15}N$ labeled histidine was adjusted to 3 mM. Each sample solution was prepared in 95% H_2O/D_2O to avoid the loss of the $C^{\epsilon 1}$ proton intensity caused by H−D exchange. The histidine sample at pH 8.2 was also used to measure the $C^{\epsilon 1}-N^{\delta 1}$ and $C^{\epsilon 1}-N^{\epsilon 2}$ coupling constants in the deprotonated form by 1-D ^{13}C-spectroscopy, and the coupling constants were in excellent agreement with published values. The ^{13}C, ^{15}N labeled N-terminal metal-binding domain of MerA (NmerA) was expressed as described previously (*14*), except that the cells were grown from the beginning in $^{13}C/^{15}N$ labeled media to maximize the level of ^{15}N incorporation.

NMR Spectroscopy. NMR experiments were measured on a Bruker 500 MHz Avance NMR instrument equipped with a triple resonance cryoprobe. The 2-D (Hβ)Cβ(CγCδ)Hδ

† This work was supported by a grant from NIH GM-56531, GM-66039, the Sandler Family Supporting Foundation (C.C. and V.D.), and NSF MLB9982576 (S.M.). N.S. was supported by Ajinomoto Co., Inc.

* To whom correspondence should be addressed. Telephone: (415) 502-7050. Fax: (415) 476-0688. E-mail: volker@picasso.ucsf.edu.

[1] Abbreviations: HMQC, heteronuclear multiple quantum coherence; HMBC, heteronuclear multiple bond coherence; HSMQC, heteronuclear single and multiple quantum coherence; NmerA, the 69 a.a. N-terminal domain of the transposon 501 (Tn501) mercuric ion reductase.

10.1021/bi0344679 CCC: $25.00
Published on Web 07/11/2003

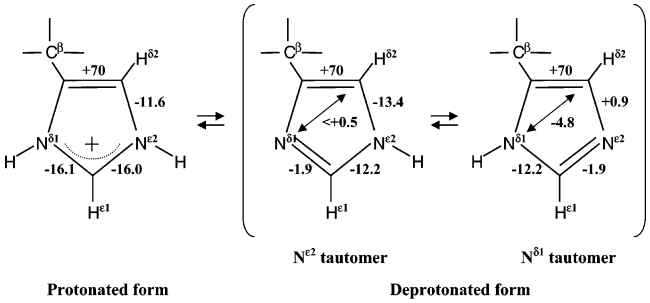

FIGURE 1: Protonated and deprotonated forms and tautomeric states of histidine residues in solution and J_{C-N}, J_{C-C} coupling constants in the imidazole ring of the histidine residue.

(*18*) and 2-D HCN (*12*) experiments were acquired as 512 × 20 and 512 × 128 complex matrixes, respectively. For the 2-D (Hβ)Cβ(CγCδ)Hδ spectrum, spectral widths of 7003 and 5031 Hz in the ^{1}H and ^{13}C dimensions were employed. The spectral widths in the ^{1}H and ^{15}N dimensions were 7003 and 10 139 Hz for the 2-D HCN spectrum.

The constant time HSQC (CT-HSQC) and ^{15}N edited CT-HQSC spectra were recorded with spectral widths of 7003 Hz for ^{1}H and 4000 Hz for ^{15}N. For the measurement of the C−N coupling constants of the free histidine at 25 °C, the ^{15}N edited CT-HSQC sequence was applied to suppress the signal from^{14}N-bound carbon coherences. CT-HSQC experiments to determine the C−N coupling constants of the histidine residues in NmerA were measured at 37 °C with the CT-HSQC sequence shown in Figure 2a. The CT-HSQC spectra of NmerA in living cells were acquired by interleaving the C−N modulated and unmodulated experiments to average potential changes of the cellular environment during the NMR measurements. The NMR data were processed with the XWINNMR software (Bruker).

RESULTS

Theoretical Basis. One-, two-, and three-bond coupling constants in the imidazole ring for the various states have been extensively studied using free histidine and 1-methylimidazole by 1-D spectroscopy (*19*, *20*) as well as fitting curves to *J*-modulated experiments (*12*). As shown in Figure 1, the values of the C−N coupling constants centered at the $C^{\epsilon 1}$ and $C^{\delta 2}$ carbons depend on the protonation and tautomeric state of the imidazole ring. As described previously, these coupling constant values can be determined from the ratio of the $C^{\epsilon 1}$-H and $C^{\delta 2}$-H cross-peaks in constant time HSQC experiments measured with and without amplitude modulation by the C−N coupling. Here, we extend the original theory to allow for a quantitative analysis of the population of the individual states. For the $C^{\epsilon 1}$-H cross-peaks, this analysis yields eq 1

$$\begin{aligned} I/I_0(C^{\epsilon 1}\text{-H}) = \cos[2\pi T\{&\alpha\beta^{N\epsilon 2}J_{C\epsilon 1-N\epsilon 2} + \\ &(1-\alpha)\beta^{N\delta 1}J_{C\epsilon 1-N\epsilon 2} + (1-\beta)^{pr}J_{C\epsilon 1-N\epsilon 2}\}] \times \\ \cos[2\pi T\{&\alpha\beta^{N\epsilon 2}J_{C\epsilon 1-N\delta 1} + (1-\alpha)\beta^{N\delta 1}J_{C\epsilon 1-N\delta 1} + \\ &(1-\beta)^{pr}J_{C\epsilon 1-N\delta 1}\}] \end{aligned} \quad (1)$$

where I and I_0 are the cross-peak intensities in the spectrum with and without the J_{C-N} amplitude modulation respectively, T is the duration of the constant time delay, α is the population of the $N^{\epsilon 2}$-H tautomer in the fraction of the deprotonated form, and β is the fraction of the deprotonated form in the total protein sample. $^{N\epsilon 2}J_{C\epsilon 1-N\epsilon 2}$, $^{N\epsilon 2}J_{C\epsilon 1-N\delta 1}$, $^{N\delta 1}J_{C\epsilon 1-N\epsilon 2}$, $^{N\delta 1}J_{C\epsilon 1-N\delta 1}$, $^{pr}J_{C\epsilon 1-N\epsilon 2}$, and $^{pr}J_{C\epsilon 1-N\delta 1}$ are the C−N coupling constants in the $N^{\epsilon 2}$-H tautomer, the $N^{\delta 1}$-H tautomer, and the protonated form, respectively.

The pK_a of the histidines can be calculated from

$$pK_a = pH - \log\{\beta/(1-\beta)\} \quad (2)$$

Combining eqs 1 and 2 yields eq 3

$$\begin{aligned} I/I_0(C^{\epsilon 1}\text{-H}) = \cos\{2\pi T[&\alpha\{1/(10^{[pK_a-pH]}+1)\}^{N\epsilon 2}J_{C\epsilon 1-N\epsilon 2} + (1-\alpha)\{1/(10^{[pK_a-pH]}+1)\}^{N\delta 1}J_{C\epsilon 1-N\epsilon 2} + \{10^{[pK_a-pH]}/(10^{[pK_a-pH]}+1)\}^{pr}J_{C\epsilon 1-N\epsilon 2}]\} \\ \cos\{2\pi T[&\alpha\{1/(10^{[pK_a-pH]}+1)\}^{N\epsilon 2}J_{C\epsilon 1-N\delta 1} + (1-\alpha)\{1/(10^{[pK_a-pH]}+1)\}^{N\delta 1}J_{C\epsilon 1-N\delta 1} + \{10^{[pK_a-pH]}/(10^{[pK_a-pH]}+1)\}^{pr}J_{C\epsilon 1-N\delta 1}]\} \end{aligned} \quad (3)$$

Note that, if $\rho \times 100\%$ is the level of ^{15}N incorporation in a histidine residue (ρ is the fraction of ^{15}N), then the intensity ratio in the CT-HSQC experiments is given by

$$\begin{aligned} I/I_0(C^{\epsilon 1}\text{-H}) = (\rho)^2\cos(\Phi)\cos(\theta) + (\rho-\rho^2)\cos(\Phi) + \\ (\rho-\rho^2)\cos(\theta) + (1-\rho)^2 \end{aligned} \quad (3')$$

where cos(Φ) and cos (θ) are the cosine functions in eq 3. The incorporation of ^{15}N is assumed to be equally distributed at the $N^{\delta 1}$ and $N^{\epsilon 2}$ positions of the histidine residues. In practice, this will become only a problem for incorporation levels of less than 95%. At this level, the maximum error of the function of I/I_0 is 0.02, which is of the same order as the error of measuring the intensities in the experiments with the free histidine. For lower incorporation levels, however, the labeling ratio can become the dominant measurement error depending on the accuracy of the intensity measurement. The error in this intensity measurement in turn depends on the individual sample since it is a function of the sensitivity of the NMR instrument as well as the nature and concentration of the sample.

The intensity ratio of the $C^{\delta 2}$-H cross-peaks, with and without the amplitude modulation, is given by

$$\begin{aligned} I/I_0(C^{\delta 2}\text{-H}) = \cos[2\pi T\{&\alpha\beta^{N\epsilon 2}J_{C\delta 2-N\epsilon 2} + (1-\alpha)\beta^{N\delta 1}J_{C\delta 2-N\epsilon 2} + (1-\beta)^{pr}J_{C\delta 2-N\epsilon 2}\}] \times \\ \cos[2\pi T\{&\alpha\beta^{N\epsilon 2}J_{C\delta 2-N\delta 1} + (1-\alpha)\beta^{N\delta 1}J_{C\delta 2-N\delta 1} + (1-\beta)^{pr}J_{C\delta 2-N\delta 1}\}] \end{aligned} \quad (4)$$

where $^{N\epsilon 2}J_{C\delta 2-N\epsilon 2}$, $^{N\epsilon 2}J_{C\delta 2-N\delta 1}$, $^{N\delta 1}J_{C\delta 2-N\epsilon 2}$, $^{N\delta 1}J_{C\delta 2-N\delta 1}$, $^{pr}J_{C\delta 2-N\epsilon 2}$, and $^{pr}J_{C\delta 2-N\delta 1}$ are again the C−N coupling constants in the $N^{\epsilon 2}$-H tautomer, the $N^{\delta 1}$-H tautomer, and the protonated form, respectively.

In the case of $\beta = 1$, since $^{N\delta 1}J_{C\delta 2-N\epsilon 2}$ (0.9 Hz) and $^{N\epsilon 2}J_{C\delta 2-N\delta 1}$ (<0.5 Hz) are negligible (Figure 1), eq 4 can be

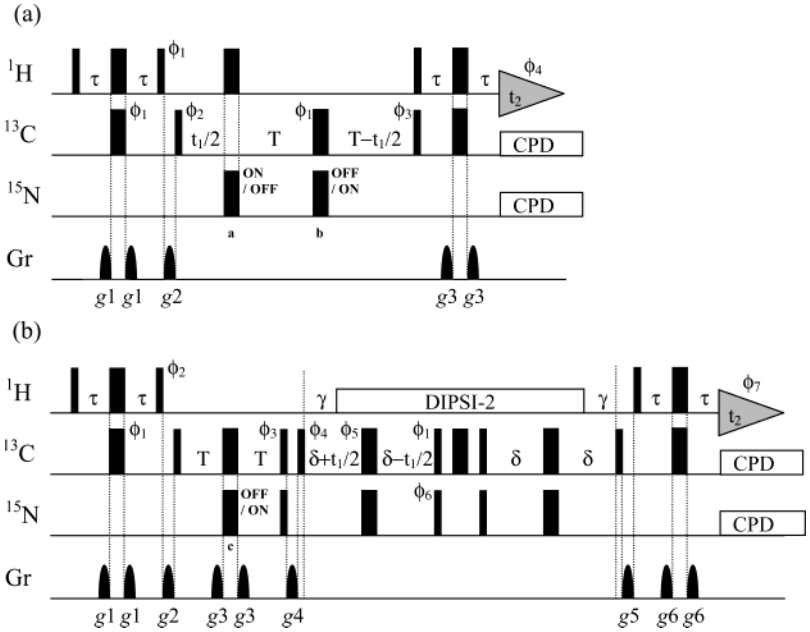

FIGURE 2: (a) Pulse sequence for CT-HSQC experiments without and with J_{C-N} amplitude modulation. Narrow and wide bars indicate 90 and 180° pulses, respectively. For measuring a spectrum without the modulation, the ^{15}N 180° pulse is applied at position a, whereas for measuring the spectrum with modulation, the ^{15}N 180° pulse is instead applied at position b. Phase cycling: $\phi_1 = y$; $\phi_2 = x$; $\phi_3 = x, -x$; and $\phi_4 = x, -x$. The phases for other pulses are fixed at x. The quadrature in F_1 was achieved via States-TPPI of ϕ_2. Delay duration: τ = 1.2 ms; gradients (sine bell shaped): $g1$ = (800 μs, 10 G/cm), $g2$ = (600 μs, 10 G/cm), and $g3$ = (600 μs, 30 G/cm). The WATERGATE water suppression scheme with a 3-9-19 refocusing pulse was incorporated into the reverse INEPT step (*22*). (b) Pulse sequence for ^{15}N edited CT-HSQC experiments, without and with J_{C-N} amplitude modulation. Many of the details of this sequence are similar to those of the CT-HSQC. For the ^{15}N edited CT-HSQC experiments without the modulation, the ^{15}N 180° pulse was not applied at position c, whereas the ^{15}N 180° pulse was applied for that with the modulation. Phase cycling: $\phi_1 = x, -x$; $\phi_2 = y, -y$; $\phi_3 = -x$; $\phi_4 = x$; $\phi_5 = 4(x), 4(y), 4(-x), 4(-y)$; $\phi_6 = 2(x), 2(-x)$; and $\phi_7 = 2(x), 4(-x), 2(x)$. The quadrature in F_1 was achieved via States-TPPI of ϕ_4. Delay duration: τ = 1.2 ms; γ = 2.4 ms; δ = 6.0 ms; and gradients (sine bell shaped): $g1$ = (600 μs, 5 G/cm), $g2$ = (800 μs, 18 G/cm), $g3$ = (1000 μs, 43 G/cm), $g4$ = (800 μs, −18 G/cm), $g5$ = (1200 μs, 28 G/cm), and $g6$ = (500 μs, 30 G/cm).

simplified to

$$I/I_0(C^{\delta 2}\text{-H}) = \cos\{2\pi T\alpha\, {}^{N\epsilon 2}J_{C\delta 2-N\epsilon 2}\} \times \cos\{2\pi T(1-\alpha)^{N\delta 1}J_{C\delta 2-N\delta 1}\} \quad (5)$$

Pulse Sequences. The method described here to measure the C−N coupling constants is a 2-D difference experiment, and the pulse sequences, which are sketched in Figure 2, are CT-HSQC type measurements (*21*). Since the details of the pulse sequence shown in Figure 2a have been discussed previously, we will here mainly focus on the improved pulse sequence shown in Figure 2b. The measured cross-peak intensities can be affected by systematic errors, such as the level of ^{15}N enrichment (*21*). This systematic error can be suppressed by the new pulse sequence shown in Figure 2b. In Figure 2b, the ^{15}N 180° pulse in the middle of the constant time period creates an antiphase coherence, which is eliminated by the subsequent gradient pulse $g4$. Without the ^{15}N 180° pulse, the magnetization is refocused at the end of the constant time period. After the constant period that either generates, or does not generate, the amplitude modulation, the magnetization is transferred from ^{13}C to ^{15}N, thereby eliminating coherences of ^{13}C spins with ^{14}N neighbors. This step ensures that only magnetization of ^{13}C spins that can be modulated by the C−N coupling contributes to the observed cross-peaks. After this filtering element, the following INEPT step refocuses the ^{13}C coherences with respect to the ^{15}N spins, and the coherences are finally transferred into in-phase proton magnetization for detection. However, a longer duration time for coherence transfers, which results from the additional INEPT steps from ^{13}C to ^{15}N and back, is required. Consequently, this pulse sequence should be applied to proteins with a low molecular weight (MW < 10 000−20 000). For larger molecular weight proteins, or proteins with a short relaxation period, the original CT-HSQC sequence is more appropriate; however, great care must be taken to maximize the incorporation of ^{15}N. Alternatively, if the percentage of ^{15}N incorporation is known and if it is randomly distributed over both the $N^{\epsilon 2}$ and the $N^{\delta 1}$ positions, eq 3′ can be used to analyze the resulting cross-peak intensities.

Adjusting the Constant Time Periods. The $C^{\epsilon 1}-N^{\epsilon 2}$ and $C^{\epsilon 1}-N^{\delta 1}$ coupling constants at the $C^{\epsilon 1}$ position for the $N^{\epsilon 2}$-H tautomer are −12.2 and −1.9 Hz, and those for the $N^{\delta 1}$-H tautomer are −1.9 and −12.2 Hz, showing that the ^{13}C magnetization in the CT-HSQC spectra for both tautomeric states of the deprotonated form is modulated by the same coupling constants. In contrast, the modulation for the protonated form is generated by the relatively large coupling constants of ${}^{pr}J_{C\epsilon 1-N\epsilon 2}$ (−16.0 Hz) and ${}^{pr}J_{C\epsilon 1-N\delta 1}$ (−16.1 Hz). To calculate the difference of [pK_a − pH] from the C−N coupling constants, the modulation at the position of $C^{\epsilon 1}$ is unquestionably the appropriate probe because of the large difference in coupling constants between the protonated and the deprotonated forms and the fact that both tautomeric states of the deprotonated form are virtually indistinguishable. On the other hand, the modulation at the $C^{\delta 2}$ position is the probe of choice to determine the ratio of the $N^{\epsilon 2}$-H and $N^{\delta 1}$-H tautomers since the ^{13}C magnetization at the $C^{\delta 2}$ position in the deprotonated form is effectively dominated by the

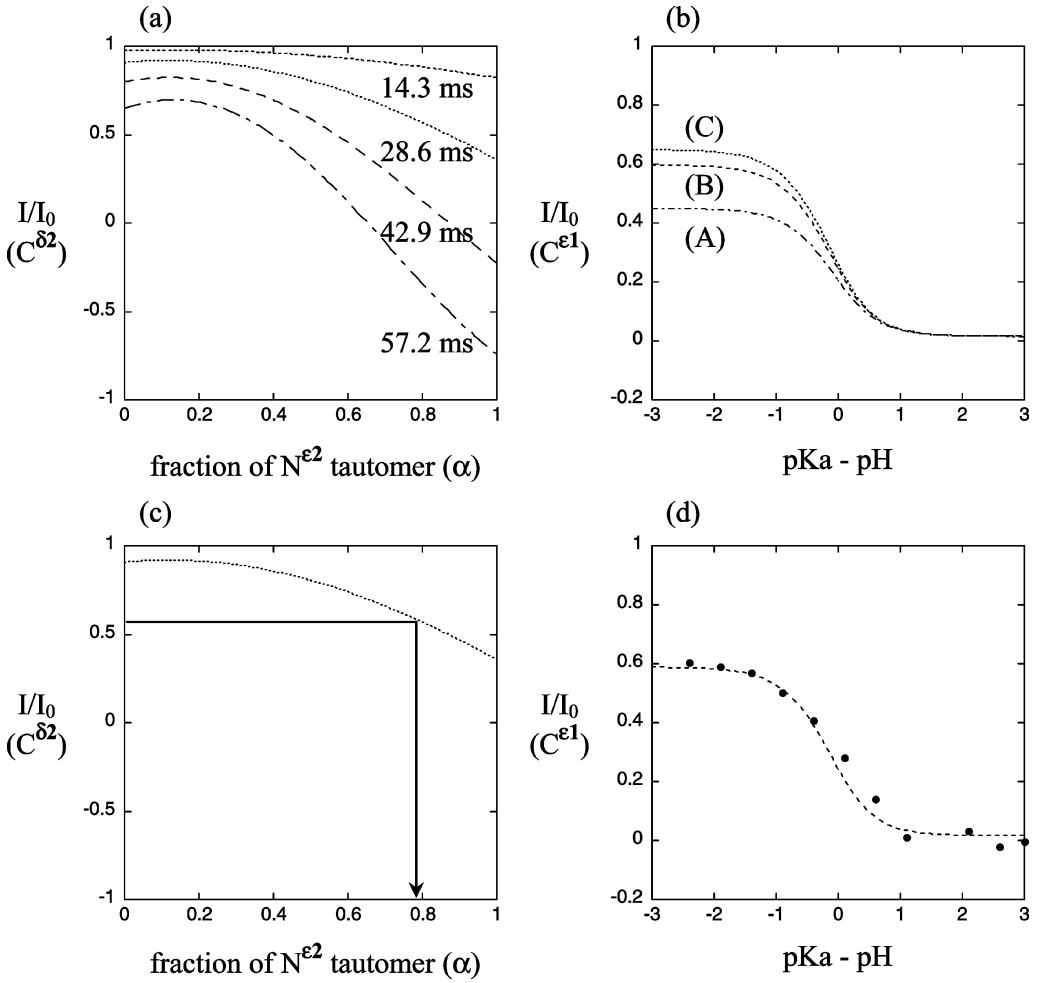

FIGURE 3: (a) Theoretical curves shown in eq 5 for the intensity ratio for the $C^{\delta 2}$-H cross-peaks vs the fraction of the $N^{\epsilon 2}$-H tautomer, using the constant times $2T = 14.3$, 28.6, 42.9, and 57.2 ms. (b) Theoretical curves shown in eq 3 for the intensity ratio for the $C^{\epsilon 1}$-H cross-peaks vs [pK_a − pH], using the ratios of the $N^{\epsilon 2}$-H and $N^{\delta 1}$-H tautomers 1:0 (A), 0.75:0.25 (B), 0.5:0.5 (C), 0.25:0.75 (B), and 0:1 (A). Note that the same curves are obtained with the inverse ratios such as 0.75:0.25 and 0.25:0.75. The constant time $2T$ was set to 28.6 ms. (c) The theoretical curve with $2T = 28.6$ ms is shown with the dashed line. The intensity ratio for the $C^{\delta 2}$-H cross-peaks in the histidine at pH 8.0 was 0.66, indicating that the fraction of the $N^{\epsilon 2}$-H tautomer is 0.77. (d) The theoretical curve with 0.77 of the fraction of the $N^{\epsilon 2}$-H tautomer is shown by the dashed line. The observed intensity ratios at several pH values are plotted with the solid circles, using the free histidine. Note that the pK_a value of the histidine was 6.3, estimated from the pH dependency of the $C^{\epsilon 1}$-H chemical shifts.

$^{N\epsilon 2}J_{C\delta 2-N\epsilon 2}$ coupling (−13.4 Hz) for the $N^{\epsilon 2}$-H tautomer and by $^{N\delta 1}J_{C\delta 2-N\delta 1}$ coupling (−4.8 Hz) for the $N^{\delta 1}$-H tautomer.

Figure 3a shows theoretical curves calculated with eq 5 for several different constant time periods $2T$. The constant time periods are set to multiples of $1/J_{C\delta 2-C\gamma}$, to eliminate the $^{13}C-^{13}C$ coupling between the $C^{\delta 2}$ and the C^{γ} carbons. As illustrated in Figure 3a, the difference in the I/I_0 ratios for the two tautomeric states becomes larger with an increase in the duration of the constant time delay $2T$. However, longer constant time delays also cause signal attenuation because of ^{13}C T_2 relaxation. For most applications to proteins a delay of 28.6 ms seems, therefore, to be the best compromise between maximizing the difference of the ratios and minimizing the signal loss due to relaxation.

Theoretical curves for the dependence of the $C^{\epsilon 1}$-H cross-peak intensities on the difference between the pK_a and the pH calculated with eq 3, in which the value of $2T$ is set to 28.6 ms, are shown in Figure 3b. Depending on the ratio between the $N^{\epsilon 2}$-H and the $N^{\delta 1}$-H tautomers, the theoretical curves differ slightly. However, the combined intensity ratio of the deprotonated states is quite different from that of the protonated state, allowing for the estimation of the ratio of both states. Therefore, the value of $2T$ was set to 28.6 ms throughout the experiments.

Verification of the Theory with Free Histidine. To verify the theory, this method was applied to histidine uniformly labeled with ^{13}C and ^{15}N. Unintentionally, ∼10% of the molecules were not completely labeled with the ^{15}N, as estimated by NMR. To suppress coherences of carbon nuclei connected to a ^{14}N nucleus, the ^{15}N edited CT-HSQC sequences shown in Figure 2b were used.

The intensity ratio of the $C^{\delta 2}$-H cross-peaks in the dominantly deprotonated form measured at pH 8.0 was 0.66, and the corresponding ratio of the $N^{\epsilon 2}$-H tautomer was 0.77 (Figure 3c). This is in very good agreement with a previous report, in which the ratio of the two tautomers was determined from peak intensities at a very low temperature where the fast chemical exchange between the two states in suppressed (*24*). That investigation had shown that under these conditions ∼80% of the histidine side chains are in the $N^{\epsilon 2}$-H tautomer. This result is in agreement with other studies that were carried out at temperatures between 25 and 30 °C (*12*, *20*, *23*).

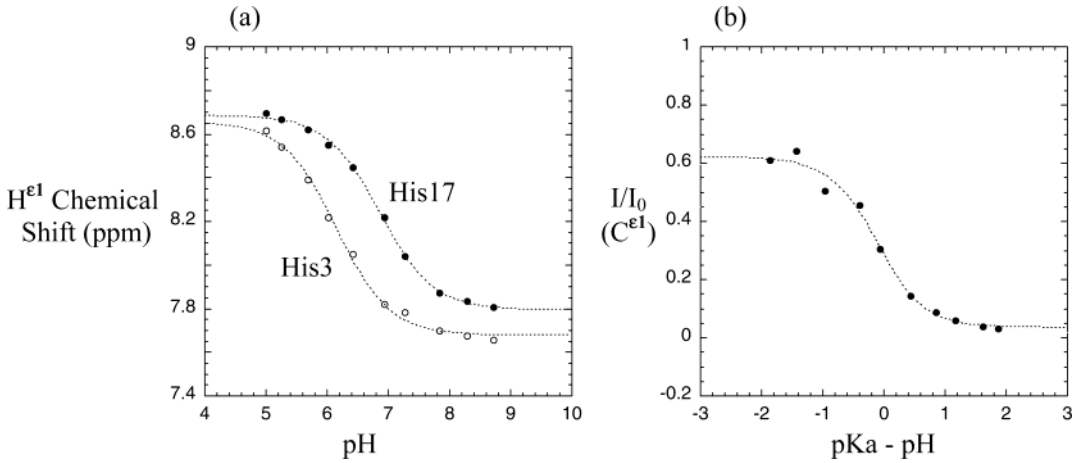

FIGURE 4: (a) pH dependencies of the $H^{\epsilon 1}$ chemical shifts for His3 and His17 in NmerA. The best-fit curves shown by the dashed lines were generated, indicating that the pK_a values of His3 and His17 were 6.20 ± 0.05 and 6.90 ± 0.05, respectively. (b) The observed intensity ratios for the $C^{\epsilon 1}$-H cross-peaks of His17 in NmerA at several pH values are plotted with the solid circles. The best-fit curve shown by the dashed line was generated from the ratio of the $N^{\epsilon 2}$-H tautomer as an unknown parameter, showing that the fraction of the $N^{\epsilon 2}$-H tautomer is 0.76 ± 0.06 or 0.24 ± 0.06.

Figure 3d shows the intensity ratios of the $C^{\epsilon 1}$-H cross-peaks at various pH values. Note that the pK_a value for free histidine is 6.3, which was verified by a pH titration experiment, and [pK_a − pH] on the *X* axis is the difference between this pK_a of the histidine and the pH of the solution. The theoretical curve is also shown in Figure 3d, with the ratio of the $N^{\epsilon 2}$-H tautomer (α in eq 3) set to 0.77, as determined from the ratio of the $C^{\delta 2}$-H cross-peaks. This result shows that the experimental data confirm the theoretical equations and demonstrate that a quantitative analysis of the intensity ratios of the $C^{\delta 2}$-H and $C^{\epsilon 1}$-H cross-peaks can yield the ratio of the tautomeric forms as well as the [pK_a − pH] difference. In other words, the pK_a or pH value can be estimated, if one of them is already known or determined by another procedure. It is, however, important to notice that the application of this method is limited to pH values in the vicinity of the pK_a of the histidine. In Figure 3d, the intensity ratio follows a sigmoidal curve as a function of the difference between pH and pK_a. At approximately 1 pH unit above and below the pK_a this curve becomes flat, and the intensity ratio does not depend on the exact pH value anymore. That means that this method can only measure pH values that differ by less than one unit from the pK_a value of the histidine. The same limitation of course also holds for measurements of the pK_a value. Most accurate results are obtained in the immediate vicinity of the pK_a where the function shows the steepest curvature.

Assignment and pK_a of Histidine Residues in NmerA. As reported previously, in-cell NMR experiments can provide information about the conformation and dynamics of proteins inside living cells (*13*−*17*). To investigate the suitability of the coupling constant method to determine the intracellular pH of bacteria, we used the N-terminal domain of the bacterial mercury-detoxification protein MerA (NmerA) as a test case. NmerA binds Hg^{2+} via its two cysteines located in a metal binding loop and the adjacent helix. Two histidine residues are present in NmerA, and one of them (His17) is located in this helix and close to these cysteine residues.

The assignments of the main chain $^1H^N$, ^{15}N, $^{13}C\alpha$, and $^{13}C\beta$ NMR signals from NmerA were derived from the 3-D HNCA, 3-D CBCACONH, and 3-D HCCCONH-TOCSY spectra. Details of the assignments will be published elsewhere. The assignments for the histidine side chain ^1H and ^{13}C signals were collected from the 2-D (Hβ)Cβ-(CγCδ)Hδ and 2-D HCN spectra. The ^1H and ^{13}C chemical shifts (ppm) at pH 6.8 for $C^{\epsilon 1}$-H and $C^{\delta 2}$-H in His3 are 7.8, 136.1 and 7.1, 117.8, respectively, and those in His17 are 8.3, 135.1 and 7.3, 117.7, respectively.

The pK_a value for His17 is 6.90 ± 0.05 as determined by the pH dependence of the $H^{\epsilon 1}$ proton chemical shift (Figure 4a). This histidine is an appropriate probe to measure the pH since the internal pH in living cells has been reported to range from pH 6.0 to 8.0 (*25*, *26*), which is within the ±1 pH unit interval from the pK_a value as discussed above. In contrast, His3, located in the N-terminus, has a pK_a value of 6.20 ± 0.05 (Figure 4a) and is, therefore, less well-suited for measuring the pH inside living cells.

pH Dependence of the Coupling Constants for NmerA. As a first step to investigate the intracellular pH of *Escherichia coli*, the pH dependence of the intensity ratios in the CT-HSQC spectra was obtained to verify the theoretical curve. It would be difficult to apply the ^{15}N edited HSQC sequence to a system of proteins in living cells, in which the large inhomogeneity leads to fast transverse relaxation. Instead, the basic CT-HSQC sequence as shown in Figure 2a was applied to the ^{13}C and ^{15}N labeled NmerA, which was carefully prepared to maximize the incorporation of ^{15}N by growing cells in $^{13}C/^{15}N$ labeled media from the beginning.

Figure 4b shows the ratio of the $C^{\epsilon 1}$-H cross-peaks of His17 measured at different pH values together with the theoretical curve calculated with eq 3. The $C^{\delta 2}$-H cross-peaks for His17 were too weak to measure the coupling constants precisely because of line broadening in the HSQC spectra. However, the fraction of the $N^{\epsilon 2}$-H tautomer for His17 was roughly estimated as 0.70−0.80, based on the intensity ratio of the weak $C^{\delta 2}$-H cross-peaks at pH 8, at which the deprotonated form is expected to be dominant. The fraction of the $N^{\epsilon 2}$-H tautomer, calculated from the best-fit curve for the experimental data, was 0.76 ± 0.06 or 0.24 ± 0.06, one of which (0.76 ± 0.06) is consistent with the value estimated from the $C^{\delta 2}$-H intensity ratio (Figure 4b). It should be noted that this best-fit curve was generated with the fraction of the $N^{\epsilon 2}$-H tautomer as an unknown parameter.

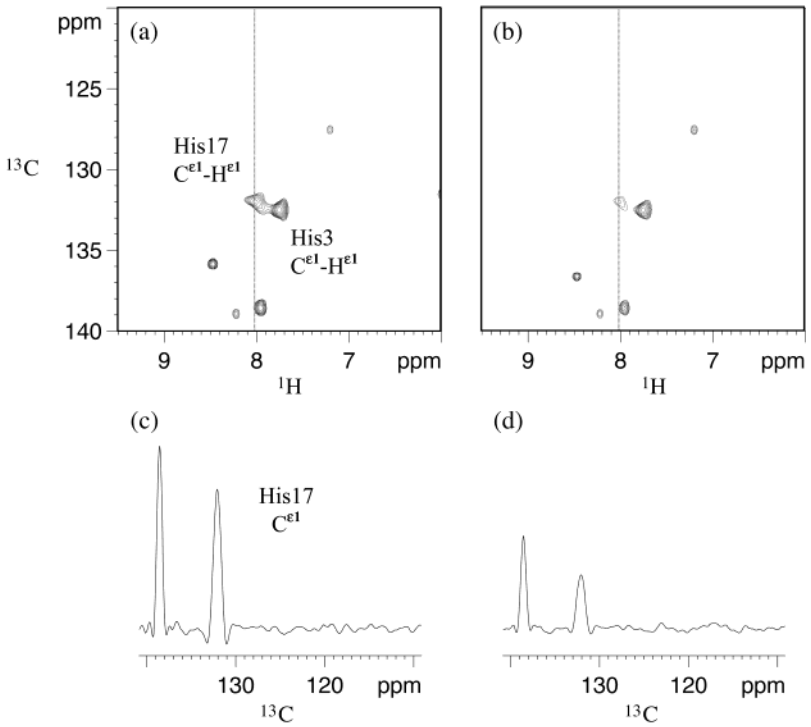

FIGURE 5: CT-HSQC spectra of uniformly ^{13}C and ^{15}N labeled NmerA in living cells. $C^{\epsilon 1}$-H regions measured without (a) and with (b) J_{C-N} amplitude modulation. (c and d) The 1-D cross-sections on top of each section were taken along the ^{13}C dimension at the positions indicated by the dashed lines.

pH Determination in Living Cells. To measure the pH in the *E. coli* cytoplasm, CT-HSQC spectra of ^{13}C and ^{15}N labeled NmerA in cells were measured without and with J_{C-N} amplitude modulation (Figure 5). The chemical shifts of the two histidine residues did not change significantly, and the intensity ratio for His17, estimated from both spectra, was 0.37 ± 0.04. Figure 5c,d shows the 1-D cross-sections, which were taken along the ^{13}C dimension at the position indicated by the dashed lines, demonstrating that the signal-to-noise ratio is reliable enough to measure the intensity ratio, even though the total measurement time for both spectra was only 30 min. On the basis of the calibration curve shown in Figure 4b, the value of [pK_a − pH] was calculated to be -0.2 ± 0.1. Combined with the pK_a of His17, the pH in the *E. coli* cytoplasm could thus be calculated to be 7.1 ± 0.1 under our experimental conditions.

DISCUSSION

The present results demonstrate that the intensity ratio determined from the CT-HSQC experiments without and with J_{C-N} amplitude modulation can be used to measure the difference of [pK_a − pH] and the ratio of $N^{\epsilon 2}$-H and $N^{\delta 1}$-H tautomers. The ^{15}N edited CT-HSQC can in principle decrease the systematic error, which is caused by the incomplete incorporation of ^{15}N within the histidine residues. However, in practice, the CT-HSQC sequences are preferable to measure the pH in cells and in applications with large molecular weight proteins because a longer duration time for coherence transfers, resulting from the additional INEPT step from ^{13}C to ^{15}N, is required for the ^{15}N edited HSQC experiments. Especially, for in-cell NMR applications, the signal-to-noise ratio would be worse in the ^{15}N edited HSQC experiments, making it difficult to measure, for example, the pH precisely. To avoid systematic errors, either the sample must be carefully prepared to maximize the ^{15}N incorporation or eq 3′ has to be used to account for the ^{14}N fraction.

Depending on the ratio of the $N^{\epsilon 2}$-H and $N^{\delta 1}$-H tautomers, the theoretical curves of I/I_0 ($C^{\epsilon 1}$-H) versus [pK_a − pH] differ slightly. In principle, however, the ratio of the $N^{\epsilon 2}$-H and $N^{\delta 1}$-H tautomers can be obtained from the intensity ratio of the $C^{\delta 2}$-H cross-peaks in the same experiment. Alternatively, the tautomeric ratio can also be obtained from the pH dependence of the intensity ratio for the $C^{\epsilon 1}$-H cross-peaks alone if, for example, the $C^{\delta 2}$-H cross-peaks cannot be detected. In these cases, measuring a calibration curve for the intensity ratio I/I_0 ($C^{\epsilon 1}$-H) versus [pK_a − pH] is a good alternative since the $C^{\epsilon 1}$-H cross-peaks are usually detected with an excellent signal-to-noise ratio because of the lack of dipolar interactions with the other protons. Using this second method that only depends on the intensity ratio of the $C^{\epsilon 1}$-H cross-peaks, we determined the intracellular pH of *E. coli* to be 7.1 ± 0.1. One prerequisite for applying this method to measure the intracellular pH, however, is that the pK_a value of the histidine does not differ between the in vivo and the in vitro condition. A change in pK_a values, however, requires a change in the conformation of the protein, for example, by forming new hydrogen bonds, which would result in differences in the chemical shifts of the histidine residues. In the case of NmerA such changes do not occur, strongly suggesting that the pK_a of His 17 is the same in the in vitro and the in vivo conditions.

The pH value of 7.1 ± 0.1 is 0.5 units lower than values published before, which determined the cytoplasmic pH in *E. coli* to be 7.6, based on the ^{14}C-labeled 5,5-dimethyl-2,4-oxazolidinedione distribution (*25*, *26*). In addition, pH values ranging from 7.3 to 7.5 have been reported, depending on the growth phase, in a ^{31}P NMR study (*27*). However, the conditions in all these studies are different from the condi-

tions used in the current investigation. In our experiments, the bacteria were induced to express a protein that changes the metabolic state of the cells. In all other studies, cells in different growth phases and without protein expression were used that can easily explain the differences in the measured values. The pH value of 7.1 was also confirmed by an independent measurement. In NmerA, the ^{15}N and $^{1}H_N$ chemical shifts of Cys11 are very sensitive to pH changes. Comparing the chemical shifts of the protein in the *E. coli* cytoplasm with a calibration curve obtained from a pH titration experiment with purified NmerA confirmed the measurement based on the histidine coupling constants.

Several other methods to measure the pH value of cells are currently available. They include NMR methods, the use of H^+-selective microelectrodes (*28*), radiolabeled membrane-permeable compounds (*25*, *26*), and methods that are based on fluorescence techniques (*29*−*34*). For NMR-based methods, both ^{31}P and ^{19}F have been used as probes in the past (*35*−*38*). The inorganic phosphate (P_i) chemical shift is pH-dependent and is therefore widely used to measure intracellular pH, but accurate measurements are difficult because the intracellular P_i is sometimes invisible or severely overlapped with the extracellular P_i (*39*). Moreover, although the ATP signals are intense, their chemical shifts are weakly dependent on pH within the intracellular range because the apparent pK_a of ATP in cells is low (*40*). As for the other peaks, their weaker intensity and their partial overlapping make it difficult to determine the pH. ^{19}F NMR has been used much less frequently, and the intracellular distribution of the fluorinated compounds used as probes has not been analyzed in detail (*41*).

While it is difficult to control the location of inorganic compounds, including ^{31}P and ^{19}F but also small organic pH-sensitive fluorophores, proteins are good intracellular pH probes since they can be targeted with the help of special peptide sequences to different destinations in the cells. This trait of proteins has been extensively used for pH measurements based on the pH dependence of the fluorescence of the green fluorescent protein (GFP) and its mutants (*33*, *34*) and has facilitated analyses of many organelles in living cells (*42*−*51*). In addition, several GFP mutants, with strong pH-dependent fluorescence, have been described (*52*−*56*). The very high sensitivity of these fluorescence-based measurements make them better suited for studying the intracellular pH than any of the NMR methods including the one described here. However, as demonstrated previously, in-cell NMR experiments provide far more detailed information about the conformation and dynamics of proteins in their natural environment. Since both conformation and dynamics—and therefore the function of proteins—depend on the protonation state, determining the intracellular pH during in-cell NMR measurements can be important. While proteins are good intracellular pH probes, because of their potential to be targeted to various organelles in the cells, not all proteins contain backbone resonances with pH-sensitive chemical shifts in the pH range found inside living cells as NmerA does. However, most proteins contain histidines, and the method described here can, therefore, be used to determine the intracellular pH with the same in-cell NMR sample used for investigating the conformation and dynamics. Fluorescence-based measurements cannot offer this advantage.

In addition to determining the intracellular pH value, the method described here is very useful to determine the pK_a values of histidine residues without pH titration in vitro. Traditionally, pH titration experiments have been used to obtain these values; however, this method cannot be applied to proteins, which are not stable over a wide pH range. Moreover, surrounding residues, whose protonation states depend on the pH, for example, another histidine, lysine, aspartic acid and so on, sometimes influence the chemical shift of the histidine residue, making it difficult to calculate its pK_a. The coupling constant method offers an attractive alternative. Moreover, a quantitative analysis of the $C^{\delta 2}$-based coupling constants also allows for the determination of the tautomeric ratio of histidine side chains.

ACKNOWLEDGMENT

The authors sincerely thank Melissa Malone for technical assistance and Anson Nomura for useful discussions.

REFERENCES

1. Tsilikounas, E., Rao, T., Gutheil, W. G., and Bachovchin, W. W. (1996) *Biochemistry 35*, 2437−2444.
2. Markley, J. L., and Westler, W. M. (1996) *Biochemistry 35*, 11092−11097.
3. Halkides, C. J., Wu, Y. Q., and Murray, C. J. (1996) *Biochemistry 35*, 15941−15948.
4. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) *Protein Sci. 2*, 543−558.
5. Bax, A., Griffey, R. H., and Hawkins, B. L. (1983) *J. Magn. Reson. 55*, 301−315.
6. Van Dijk, A. A., Scheek, R. M., Dijkstra, K., Wolters, G. K., and Robillard, G. T. (1992) *Biochemistry 31*, 9063−9072.
7. Bax, A., and Summers, M. F. J. (1986) *J. Am. Chem. Soc. 108*, 2093−2094.
8. Schmidt, J. M., Thuring, H., Werner, A., Ruterjans, H., Quaas, R., and Hahn, U. (1991) *Eur. J. Biochem. 197*, 643−653.
9. Zuiderweg, E. R. P. (1990) *J. Magn. Reson. 86*, 346−357.
10. Xia, B., Cheng, H., Skjeldahl, L., Coghlan, V. M., Vickery, L. E., and Markley, J. L. (1995) *Biochemistry 34*, 180−187.
11. Shimba, N., Takahashi, H., Sakakura, M., Fujii, I., and Shimada, I. (1998) *J. Am. Chem. Soc. 120*, 10988−10989.
12. Sudmeier, J. L., Ash, E. L., Günther, U. L., Luo, X., Bullock, P. A., and Bachovchin, W. W. (1996) *J. Magn. Reson. Ser. B 113*, 236−247.
13. Serber, Z., Keatinge-Clay, A. T., Ledwidge, R., Kelly, A. E., Miller, S. M., and Dötsch, V. (2001) *J. Am. Chem. Soc. 123*, 2446−2447.
14. Serber, Z., Ledwidge, R., Miller, S. M., and Dötsch, V. (2001) *J. Am. Chem. Soc. 123*, 8895−8901.
15. Serber, Z., and Dötsch, V. (2001) *Biochemistry 40*, 14317−14323.
16. Wieruszeski, J. M., Bohin, A., Bohin, J. P., and Lippens, G. (2001) *J. Magn. Reson. 151*, 118−123.
17. Dedmon, M. M., Patel, C. N., Young, G. B., and Pielak, G. J. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 12681−12684.
18. Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) *J. Am. Chem. Soc. 115*, 11054−11055.
19. Blomberg, F., Maurer, W., and Ruterjans, H. (1977) *J. Am. Chem. Soc. 99*, 8149−8159.
20. Alei, M., Jr., Morgan, L. O., Wageman, W. E., and Whaley, T. W. (1980) *J. Am. Chem. Soc. 102*, 2881−2887.
21. Vuister, G. W., Wang, A. C., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 5334−5335.
22. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson. Ser. A 102*, 241−245.
23. Reynolds, W. F., Peat, I. R., Freedman, M. H., and Lyerla, J. R., Jr. (1973) *J. Am. Chem. Soc. 95*, 328−331.
24. Farr-Jones, S., Wong, W. Y. L., Gutheil, W. G., and Bachovchin, W. W. (1993) *J. Am. Chem. Soc. 115*, 6813−6819.
25. Padan, E., Zilberstein, D., and Rottenberg, H. (1976) *Eur. J. Biochem. 63*, 533−541.
26. Padan, E., and Schuldiner, S. (1987) *J. Membr. Biol. 95*, 189−198.

27. Noguchi, Y., Shimba, N., Toyosaki, H., Ebisawa, K., Kawahara, Y., Suzuki, E., and Sugimoto, S. (2002) *J. Microbiol. Methods 51*, 73−82.
28. Gerson, D. F., and Burton, A. C. (1977) *J. Cell Physiol. 91*, 297−303.
29. Grinstein, S., Cohen, S., and Rothstein, A. (1984) *J. Gen. Physiol. 83*, 341−369.
30. Gillies, R. J., Martinez-Zaguilan, R., Martinez, G. M., Serrano, R., and Perona, R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7414−7418.
31. Haworth, R. S., and Fliegel, L. (1993) *Mol. Cell Biochem. 124*, 131−140.
32. Pena, A., Ramirez, J., Rosas, G., and Calahorra, M. (1995) *J. Bacteriol. 177*, 1017−1022.
33. Kneen, M., Farinas, J., Li, Y., and Verkman, A. S. (1998) *Biophys. J. 74*, 1591−1599.
34. Llopis, J., McCaffery, J. M., Miyawaki, A., Farquhar, M. G., and Tsien, R. Y. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 6803−6808.
35. Navon, G., Shulman, R. G., Yamane, T., Eccleshall, T. R., Lam, K. B., Baronofsky, J. J., and Marmur, J. (1979) *Biochemistry 18*, 4487−4499.
36. Barton, J. K., Den Hollander, J. A., Lee, T. M., MacLaughlin, A., and Shulman, R. G. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2470−2473.
37. Gillies, R. J., Ugurbil, K., Den Hollander, J. A., and Shulman, R. G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2125−2129.
38. Civan, M. M., Williams, S. R., Gadian, D. G., and Rozengurt, E. (1986) *J. Membr. Biol. 94*, 55−64.
39. Avison, M. J., Hetherington, H. P., and Shulman, R. G. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 377−402.
40. Pollard, H. B., Shindo, H., Creutz, C. E., Pazoles, C. J., and Cohen, J. S. (1979) *J. Biol. Chem. 254*, 1170−1177.
41. Taylor, J., and Deutsch, C. (1998) *Biophys. J. 53*, 227−233.
42. Lim, C. R., Kimata, Y., Oka, M., Nomaguchi, K., and Kohno, K. (1995) *J. Biochem. 118*, 13−17.
43. Zolotukhin, S., Potter, M., Hauswirth, W. W., Guy, J., and Muzyczka, N. (1996) *J. Virol. 70*, 4646−4654.
44. Hampton, R. Y., Koning, A., Wright, R., and Rine, J. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 828−833.
45. Rizzuto, R., Brini, M., Pizzo, P., Murgia, M., and Pozzan, T. (1995) *Curr. Biol. 5*, 635−642.
46. Rizzuto, R., Brini, M., De Giorgi, F., Rossi, R., Heim, R., Tsien, R. Y., and Pozzan, T. (1996) *Curr. Biol. 6*, 183−188.
47. De Giorgi, F., Brini, M., Bastianutto, C., Marsault, R., Montero, M., Pizzo, P., Rossi, R., and Rizzutto, R. (1996) *Gene 173*, 113−117.
48. Girotti, M., and Banting, G. (1996) *J. Cell Sci. 109*, 2915−2926.
49. Terasaki, M., Jaffe, L. A., Hunnicutt, G. R., and Hammer, J. A. (1996) *Dev. Biol. 179*, 320−328.
50. Cole, N. B., Smith, C. L., Sciaky, N., Terasaki, M., Edidin, M., and Lippincott-Schwartz, J. (1996) *Science 273*, 797−801.
51. Liu, J., Hughes, T. E., and Sessa, W. C. (1997) *J. Cell Biol. 137*, 1525−1535.
52. Heim, R., Prasher, D. C., and Tsien, R. Y. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 12501−12504.
53. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) *Nature 373*, 663−664.
54. Anderson, M. T., Tjioe, I. M., Lorincz, M. C., Parks, D. R., Herzenberg, L. A., Nolan, G. P., and Herzenberg, L. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8508−8511.
55. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) *Gene 173*, 33−38.
56. Kimata, Y., Iwaki, M., Lim, C. R., and Kohno, K. (1997) *Biochem. Biophys. Res. Commun. 232*, 69−73.

BI0344679