J Biomol NMR (2012) 54:199–209
DOI 10.1007/s10858-012-9664-z

ARTICLE

Lysine methylation strategies for characterizing protein conformations by NMR

Sacha Thierry Larda · Michael P. Bokoch · Ferenc Evanics · R. Scott Prosser

Received: 23 July 2012 / Accepted: 17 August 2012 / Published online: 8 September 2012

Abstract In the presence of formaldehyde and a mild reducing agent, reductive methylation is known to achieve near complete dimethylation of protein amino groups under non-denaturing conditions, in aqueous media. Amino methylation of proteins is employed in mass spectrometric, crystallographic, and NMR studies. Where biosynthetic labeling is prohibitive, amino ^{13}C-methylation provides an attractive option for monitoring folding, kinetics, protein–protein and protein-DNA interactions by NMR. Here, we demonstrate two improvements over traditional ^{13}C-reductive methylation schemes: (1) By judicious choice of stoichiometry and pH, ε-aminos can be preferentially monomethylated. Monomethyl tags are less perturbing and generally exhibit improved resolution over dimethyllysines, and (2) By use of deuterated reducing agents and ^{13}C-formaldehyde, amino groups can be labeled with $^{13}CH_2D$ tags. Use of deutero-^{13}C-formaldehyde affords either $^{13}CHD_2$, or $^{13}CD_3$ probes depending on choice of reducing agent. Making use of ^{13}C–^{2}H scalar couplings, we demonstrate a filtering scheme that eliminates natural abundance ^{13}C signal.

Keywords Reductive methylation · Lysine · NMR · Lysozyme · Tagging · Bioconjugation

Introduction

NMR studies of proteins generally require ^{2}H, ^{13}C, and ^{15}N isotopic enrichment or, at the very least, incorporation of precursors for ^{13}C-methyl labeling (Sprangers et al. 2007). Isotope labeling is necessary to accomplish assignments and perform detailed analyses of structure and dynamics. Although enrichment through minimal media is easily achieved with proteins expressed in *Escherichia coli*, many proteins require eukaryotic expression systems to provide the proper protein folding machinery and accomplish the requisite post-translational modifications. Under such circumstances, biosynthetic isotopic enrichment may be unfeasible. As an alternative, ^{13}C-methyl tagging of lysine residues provides domain specific information on conformation and dynamics. NMR applications have primarily made use of cysteine- (Kalbitzer et al. 1992; Klein-Seetharaman et al. 1999; Luchette et al. 2002; Oxenoid et al. 2002) and lysine-specific tags (Means and Feeney 1968, Zhang and Vogel 1993; Macnaughtan et al. 2005; Gerken et al. 1982), although other residues may be tagged under aqueous, non-denaturing conditions (Means and Feeney 1968; Joshi et al. 2004; Antos and Francis 2006; Ojida et al. 2005; Means and Feeney 1995; Abe et al. 1997).

Electronic supplementary material The online version of this article (doi:10.1007/s10858-012-9664-z) contains supplementary material, which is available to authorized users.

S. T. Larda · F. Evanics · R. S. Prosser (✉)
Department of Chemical and Physical Sciences, University of Toronto, UTM, 3359 Mississauga Rd. North, Mississauga, ON L5L 1C6, Canada
e-mail: scott.prosser@utoronto.ca

S. T. Larda
e-mail: s.larda@utoronto.ca

F. Evanics
e-mail: fevanics@gmail.com

M. P. Bokoch
Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA 94143, USA
e-mail: bokochm@anesthesia.ucsf.edu

R. S. Prosser
Department of Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada

Lysines are particularly convenient labeling sites due in part to their abundance in proteins of 6–7 % (Yokoyama 2003). They are often found in structurally and functionally important regions including binding interfaces and salt linkages. Both N-terminal α-NH_2 groups and lysine ε-NH_2 groups may be easily labeled under non-denaturing conditions through reductive methylation using ^{13}C-formaldehyde and a suitable reducing agent. The resulting methyl probes can be resolved in a ^{13}C,^{1}H HSQC spectrum, allowing the study of structure and dynamics of specific protein domains. Assignments are made possible by preparing lysine to arginine substitution mutants, should more traditional NMR assignment approaches fail. A significant body of research has been published on the use of ^{13}C-dimethyl lysine probes for NMR studies of protein conformational changes through intermolecular interactions, metal or small molecule binding, or changing physiological conditions (Gerken et al. 1982; Abraham et al. 2009; White and Rayment 1993).

Lysine methylation is generally non-perturbing because only a small exogenous chemical moiety is added to the native protein and charge is maintained, unlike acylation. The majority of spectroscopic and crystallographic studies have found that lysine dimethylation results in relatively minor structure perturbations (Zhang et al. 1994; Moore et al. 1998; Rayment 1997; Rypniewski et al. 1993; White and Rayment 1993), although crystallization properties of proteins may be significantly altered (Rayment 1997). Hen egg white lysozyme has long served as a model system for protein biochemistry and structural biology. X-ray crystallography studies of fully dimethylated HEWL reveal that the protein backbone is superimposable with that of the unmethylated protein to an RMSD of 0.4 Å, with the exception of two surface loops (Palmer et al. 2001). The effect of reductive methylation on protein activity is more case specific. For example, dimethylation slows the steady-state rate of MgATP hydrolysis under saturating actin concentrations for the 20 kDa myosin subfragment-1 enzyme while the catalytic activities of trypsin and chymotrypsin are unchanged (Korzhnev et al. 2005). Conversely, reductively methylated horse liver alcohol dehydrogenase was reported to have enhanced catalytic activity (Tsai et al. 1974) whereas dimethylation of the β_2 adrenergic receptor was shown to produce no significant alteration of the GPCR's structure, ligand binding or G-protein coupling relative to the unlabeled receptor (Bokoch et al. 2010).

^{13}C-methyl lysine probes generally serve as excellent NMR reporters due to favorable chemical shift dispersions, relatively long transverse relaxation times, and high sensitivity (Abraham et al. 2009; White and Rayment 1993). One drawback, however, is that dimethyl resonances are often line broadened by intermediate or slow conformational exchange between unique methyl environments (Gerken et al. 1982; Abraham et al. 2009). This exchange broadening in combination with resonance proximity/overlap with other CH-containing species in solution makes interpretation of ^{13}C,^{1}H spectra difficult, especially for very large systems.

We describe a methylation protocol that favors monoalkylation of protein amino groups and demonstrate that monomethyl tags circumvent problems typically associated with exchange broadened dimethylaminos. We also discuss modifications to reductive alkylation protocols which result in the production of $^{13}CH_2D$, $^{13}CHD_2$, or $^{13}CD_3$ tags. The presence of deuterium permits the use of NMR pulse sequence schemes that serve to filter out natural abundance ^{13}C signal, thereby making ^{13}C,H spectroscopy a viable option in situations where the labeled fraction is very small. Furthermore, deuterium is an excellent probe for the investigation of dynamics by NMR (Tugarinov and Kay 2006).

Materials and methods

Lysozyme from hen egg white (HEWL), bovine serum albumin (BSA), sodium cyanoborohydride, and dimethylamine (DMA) borane were obtained from Sigma-Aldrich. Aqueous ^{13}C formaldehyde (20 % w/w) was purchased from Cambridge Isotope Laboratories Inc. (Andover, MA). *Xenopus* calmodulin was expressed in *E. coli* and purified according to previously described methods (Ikura et al. 1990). Polyvinylpyrrolidone (PVP-40) was purchased from Sigma-Aldrich.

NMR

All NMR experiments were performed on a 600 MHz Varian Inova spectrometer using an HCN/FCN cryogenic probe. Typical ^{13}C and ^{1}H pulse lengths were 14 and 8.5 μs, respectively. Standard ^{13}C,^{1}H HSQC spectra were obtained using 8 scans and 128 increments. A spectral width of 15,000 Hz was used in the indirect (^{13}C) dimension with WALTZ-16 ^{13}C-decoupling. For ^{2}H-edited ^{13}C,^{1}H HSQC spectra of apo-calmodulin in 300 mg/mL PVP, standard pulse widths for ^{1}H, ^{13}C, and ^{2}H were 8.2, 13.8, and 515 μs, respectively. Spectra in PVP were acquired with 64 scans and 96 increments at a temperature of 20 °C. A spectral width of 6,000 Hz was used in the indirect dimension. Spectral analyses were performed using NMRPipe (Delaglio et al. 1995) and NMRViewJ (Johnson and Blevins 1994). All sample volumes were between 500 and 550 μL and contained 5–10 % D_2O. Assignments of lysozyme methyl lysine spectra were determined by corroborating those of (Macnaughtan et al. 2005) and by comparing aromatic NOEs and line broadening from dissolved nitroxide spin-labels (4-Hydroxy-TEMPO) to the predicted effects from the X-ray

crystal structure. The N-terminal amino group resonances of both the mono and dimethyl species were confirmed by: (1) the unique pH dependence of the chemical shift and (2) the pronounced intensity of the dimethyl species and relatively weak intensity of the monomethyl species of K1α in agreement with its reactivity across a range of formaldehyde concentrations.

Typical methylation reaction

HEWL, BSA, or apo-calmodulin was dissolved in 0.05–0.1 M phosphate, pH 7.5 to a concentration of 0.3–1.5 mM. ^{13}C formalin (aqueous formaldehyde) was added directly to the protein solution at a concentration equal to that of the reactive amines (N-terminal residue plus lysines). The solution was then briefly mixed and left to sit at room temperature for 1 h without stirring. Solid sodium cyanoborohydride (threefold to fivefold excess over total aldehyde) was then added directly to the mixture followed by several inversions to mix. The sample was left at room temperature for 4–6 h, followed by dialysis or exchange into appropriate buffer using centrifugal concentrators. This scheme gives roughly a 1:1 ratio of mono:dimethyl lysine.

Dimethyl-selective reaction

Lyophilized protein (HEWL or BSA) was dissolved in 0.1 M phosphate pH 7.0 to a concentration of 0.3–1.3 mM. ^{13}C formalin was added at a ratio of five equivalents aldehyde relative to the number of reactive amines. The reaction mixture was allowed to equilibrate for a minimum of 1 h at room temperature before addition of solid sodium cyanoborohydride ($\sim$2 equivalents relative to aldehyde). After 6 h at room temperature, the above scheme was repeated (for a total of 10 equivalents of ^{13}C formaldehyde).

Monomethyl-selective reaction

Lyophilized protein was dissolved in 0.1 M phosphate pH 5.7 to a concentration of 1.3 mM. ^{13}C-formalin was added to achieve equimolar stoichiometry between aldehyde and the number of reactive amines. The mixture was left to equilibrate for at least 1 h at room temperature before addition of roughly 15 equivalents of solid sodium cyanoborohydride. Reactions were left to proceed for a minimum of 6 h prior to extensive dialysis.

Aldehyde stoichiometry reaction series

HEWL

Lyophilized HEWL was dissolved to a concentration of 2 mM in 20 mM HEPES, pH 7.5. Separate reactions were conducted with the following ratios of ^{13}C-formaldehyde to reactive primary amine: 0.2:1, 0.5:1, 1:1, 2:1, 5:1, and 10:1. Each sample was reduced with 100 mM DMA-borane. After 12 h reacting at 20 °C, samples were dialyzed extensively into 20 mM HEPES, pH 7.5 and protein was then concentrated to $\sim$0.5 mM prior to NMR experiments.

BSA

A stock solution of 300 μM BSA in 50 mM phosphate, pH 6.0 was prepared. Overnight reactions were conducted at 23 °C, each with one of following aldehyde:reactive amine stoichiometries: 0.1:1, 0.25:1, 0.5:1, 1:1, 2:1, and 5:1. Samples were then extensively dialyzed back into 50 mM phosphate, pH 6.0. Final concentration of protein in each sample was 250 μM.

Lineshape estimation for exchange-broadened resonances

We used the non-iterative MEXICO program (MEXICO software suite v.3, Alex Bain, McMaster University) to predict line shapes for exchange-broadened mono and dimethyl-lysines. Deprotonation rates were calculated according to previously established methods (Borisenko et al. 2000) and populations of protonated and deprotonated species were determined from pKas determined by pH titration of hewl mono- and dimethyl lysines (Online Resource 1-7).

Deuteromethylation of ^{15}N-CaM

400 μM apo-^{15}N-calmodulin in 50 mM phosphate buffer, pH 7.5 was reacted with ^{13}C formalin at a stoichiometry of 3 aldehydes per reactive amine. The mixture was left at room temperature for 1 h before mixing with 5 molar equivalents of sodium cyanoborodeuteride relative to aldehyde. The reaction was left to proceed at room temperature for 6 h before repeating the entire sequence a second time. In total, 6 molar equivalents of ^{13}C-formaldehyde and 10 equivalents of sodium cyanoborodeuteride, relative to moles of reactive amine, were added to the solution of protein. The solution was purified by dialysis into 50 mM phosphate, pH 7.5.

Calmodulin in PVP

A 400 μM solution of partial mono/di-$^{13}CH_2D$ deuteromethylated calmodulin in phosphate was added to solid PVP. The mixture was gently shaken until the solid was fully solvated. 400 μL of the viscous mixture was combined with 50 μL D_2O and loaded into a salt tolerant susceptibility matched slot shaped Shigemi tube. The final concentration of PVP was 300 mg/mL.

Results and Discussion

A comparison of mono- and dimethyl lysine spectroscopic features

Using a typical reductive methylation scheme, described in the "Materials and methods" section, hen egg white lysozyme was labeled at pH 7.4 so as to produce a roughly equivalent quantity of monomethyl and dimethyl lysines.

Figure 1 shows the mono and dimethyl regions and assignments for each of the methyl resonances. Note that K1(ε) residue is represented by two distinct peaks at physiological pH and temperature in the dimethyl and monomethyl regions. This is consistent with the X-ray crystal structure of the lysine dimethylated protein, which reveals that the K1 ε-NH_3 is involved in a salt linkage with E7 (Kumar and Nussinov 1999; Kumar et al. 2000; Sarakatsannis and Duan 2005). A similar dimethyl lysine resonance splitting is observed for K13, which is thought to be involved in a salt link with the carboxy terminal residue, Leu129 (Imoto et al. 1972). Salt linkages presumably slow the otherwise rapid interconversion of non-degenerate methyl conformers, such that each state of the K1(ε) and K13 dimethyl aminos gives rise to a distinct resonance.

Figure 2 compares the ^{13}C,^{1}H HSQC spectra of both mono- and dimethyl lysines as a function of temperature. Resonances associated with the dimethyl lysines generally exhibit a significant degree of line broadening between 0 and 40 °C. In particular, resonance K1(ε) is represented by one motionally averaged resonance at 55 °C and two distinct resonances at 0 °C, as is also the case for K13. In contrast, the monomethyl resonances are generally well-resolved over the same temperature range.

Of note is that monomethylation (specifically, partial mono/dimethylation) does not appear to give rise to inhomogeneous broadening, which might be expected for a superposition of protein conformers with differing methylation levels. This further suggests that methylation is minimally perturbing to the structure of HEWL, although we do find that monomethylation of the protein leads to reduced solubility at pH 8.5–9.0 relative to the unmethylated or fully dimethylated form; an observation that has been corroborated by others (Gerken et al. 1982).

While the broadening of the K1(ε) and K13 resonances can be explained in part by their participation in salt-linkages (with breakage and reformation giving rise to two unique environments), we consider below, possible sources for broadening of non salt-linked dimethyl species. These include: (1) hydrogen bonding of the amino group to other residues or solvent (Moult et al. 1976), (2) cation-π interactions (3) contacts between the amino-methyl moiety and hydrophobic regions of the protein, and (4) deprotonation-controlled nitrogen inversions which reorient methyl

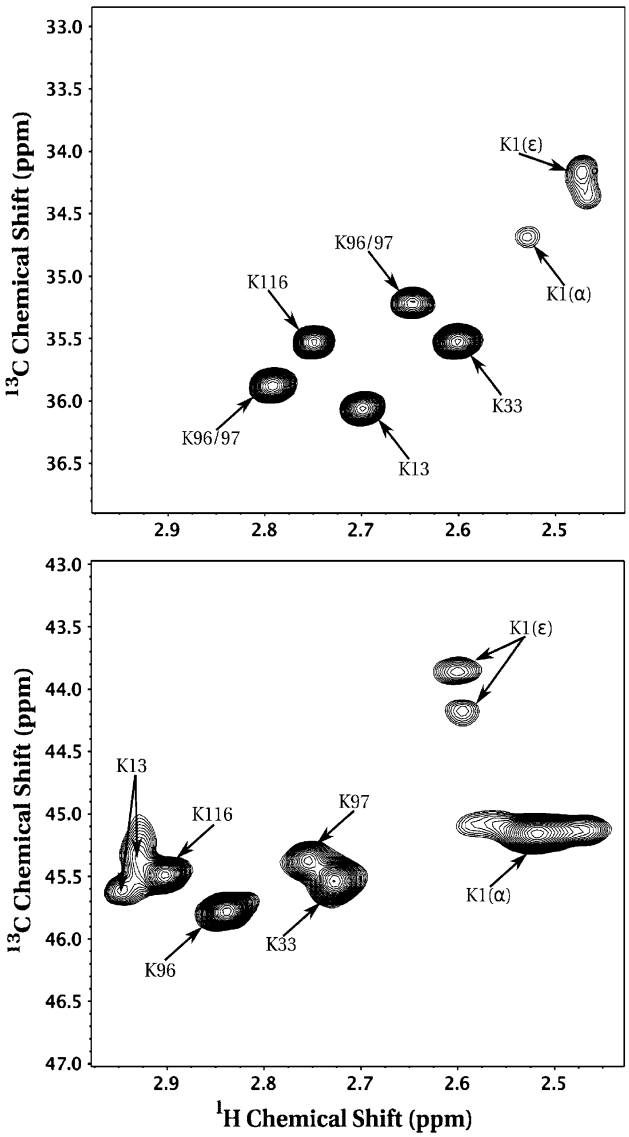

Fig. 1 (^{13}C,^{1}H) HSQC spectra at 37 °C and pH 7.5 of partially methylated HEWL showing assignments for mono- (*top*) and dimethyl-lysine resonances (*bottom*)

groups between distinct, local chemically inequivalent sites (Goux et al. 1984).

The stark difference in resolution between mono- and dimethylamino resonances can be explained by the fact that the above processes are fast on the NMR timescale for the mono species, but sufficiently slow for dimethyls so as to give rise to exchange broadening. To investigate the influence of protonation-deprotonation of the ε-amino groups on methyl dynamics, ^{13}C,^{1}H HSQC spectra were recorded as a function of pH for both mono- and dimethyl species. Increasing pH dramatically reduces line broadening of dimethyl resonances (Abraham et al. 2009) as shown by the pH-dependent spectra in Fig. 3. In contrast, pH has a small effect on the line widths associated with monomethyl lysine resonances, whose

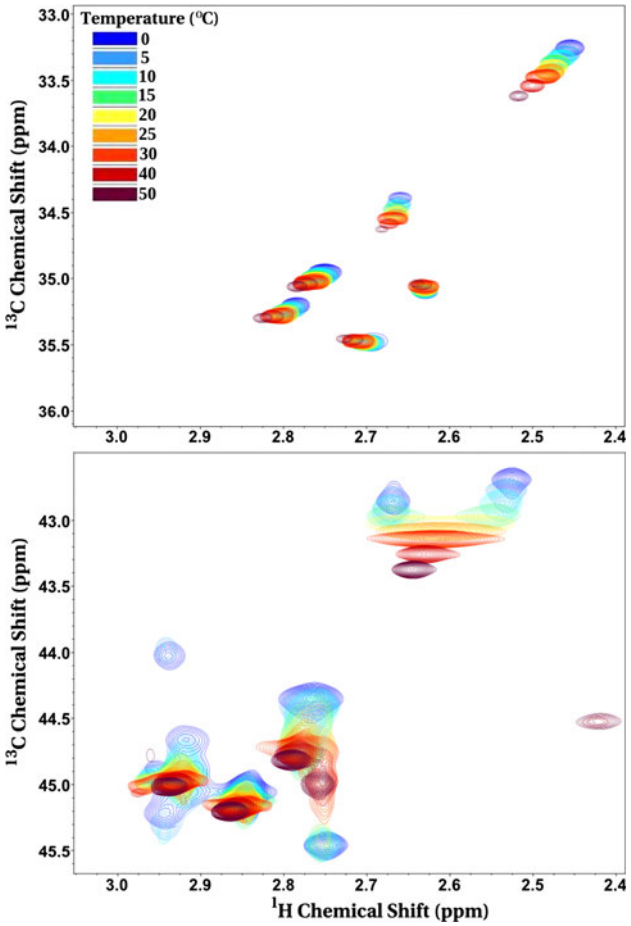

Fig. 2 (^{13}C,^{1}H) HSQC spectra of reductively methylated HEWL, as a function of temperature at pH 7.4. The dimethyl lysine corresponding to K1(ε) exists as two distinct resonances at lower temperatures as is the case for K13; many resonances exhibit coalescence and line broadening at physiological temperatures. The latter two residues are implicated in salt linkages to E7 and L129, respectively. Note that the ^{1}H and ^{13}C referencing have been altered to maximize overlap of peaks in order to emphasize the changes in linewidth that occur with temperature

reorientations are likely in the fast motional limit in the majority of cases. In addition, the size of the dimethyl moieties exacerbates local steric crowding effects whereas the smaller monomethyl lysine tags, experience greater freedom with regard to N–C bond isomerizations and nitrogen inversions. Our observations suggest that monomethyl lysine reorientations are sufficiently fast that the variety of environments sampled average to a single isotropic state, whereas the slow reorientations accompanying dimethyl lysines result in sufficiently long lived sampling of local heterogeneity (Abraham et al. 2009; Gerken et al. 1982).

Lysine side chain dynamics and deprotonation rates

An analysis of the line broadening at 25 °C, reveals that the dimethyls of K1(ε) undergo two-site exchange between similarly populated states with a correlation time of roughly 1.2 ms ($k_{ex} = 10^{1}$ s^{-1} at 25 °C, 2×10^{2} s^{-1} at 40 °C). A similar degree of broadening is observed for K13. In general, slow lysine side chain reorientations likely depend in part on the charge state of the ε-amino group since an increase in pH to within roughly 1.5 pH units of the pKa results in dramatically improved resolution. Transient deprotonation events are expected to be very fast near the dimethyl lysine pKa, as discussed by others (Borisenko et al. 2000; Tieleman et al. 2003). Lysine protonation/deprotonation kinetics are governed by the pKa of the lysine residue and the buffer, B (Borisenko et al. 2000).

Upon deprotonation, salt linkages and interactions with negatively charged sites on the protein surface become weakened allowing for more frequent reorientations of lysine side chains. Furthermore, deprotonation by either solvent or adjacent titrating species in a protein is a critical step preceding nitrogen inversion (Goux et al. 1984). Deprotonation kinetics are well known to occur through interactions with water or the buffer base as described by the following reactions:

$$\mathrm{Lys-NH_3^+ + OH^- \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} Lys-NH_2 + H_2O} \quad (1)$$

$$\mathrm{Lys-NH_3^+ + H_2O \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} Lys-NH_2 + H_3O^+} \quad (2)$$

$$\mathrm{Lys-NH_3^+ + B \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} Lys-NH_2 + BH^+} \quad (3)$$

Following the approach of Borisenko et al. (2000) we assume proton transfer to be diffusion-controlled in the thermodynamically favorable direction with a rate constant $k = 10^{10}$ M^{-1} s^{-1}, and smaller by a factor ΔpKa (i.e. the difference in the pKa values of the partners exchanging a proton) in the reverse direction. It then becomes possible to evaluate the average deprotonation rate associated with Lys-NH_3^+ for a given pKa. Note that the dimethyl lysines of lysozyme possess an average pKa around 10.5 while the average pKa of monomethyl lysines is estimated to be 11.9 (Online Resource 1). Figure 4 graphs the estimated molecular deprotonation rates for mono- and dimethyl lysines, based on the sum of the above three mechanisms, and assuming a buffer concentration of 50 mM and buffer pKa of 7.2. Exchange with either hydroxide ions or the buffer tend to make the greatest contribution to the overall deprotonation rate, which is predicted to be on an intermediate NMR timescale at physiological pH (*vida infra*).

Using the predicted kinetic rates and weighted pH-dependent populations of protonated and deprotonated species, line shape simulations could not account for the exchange broadening observed for the dimethyl lysine

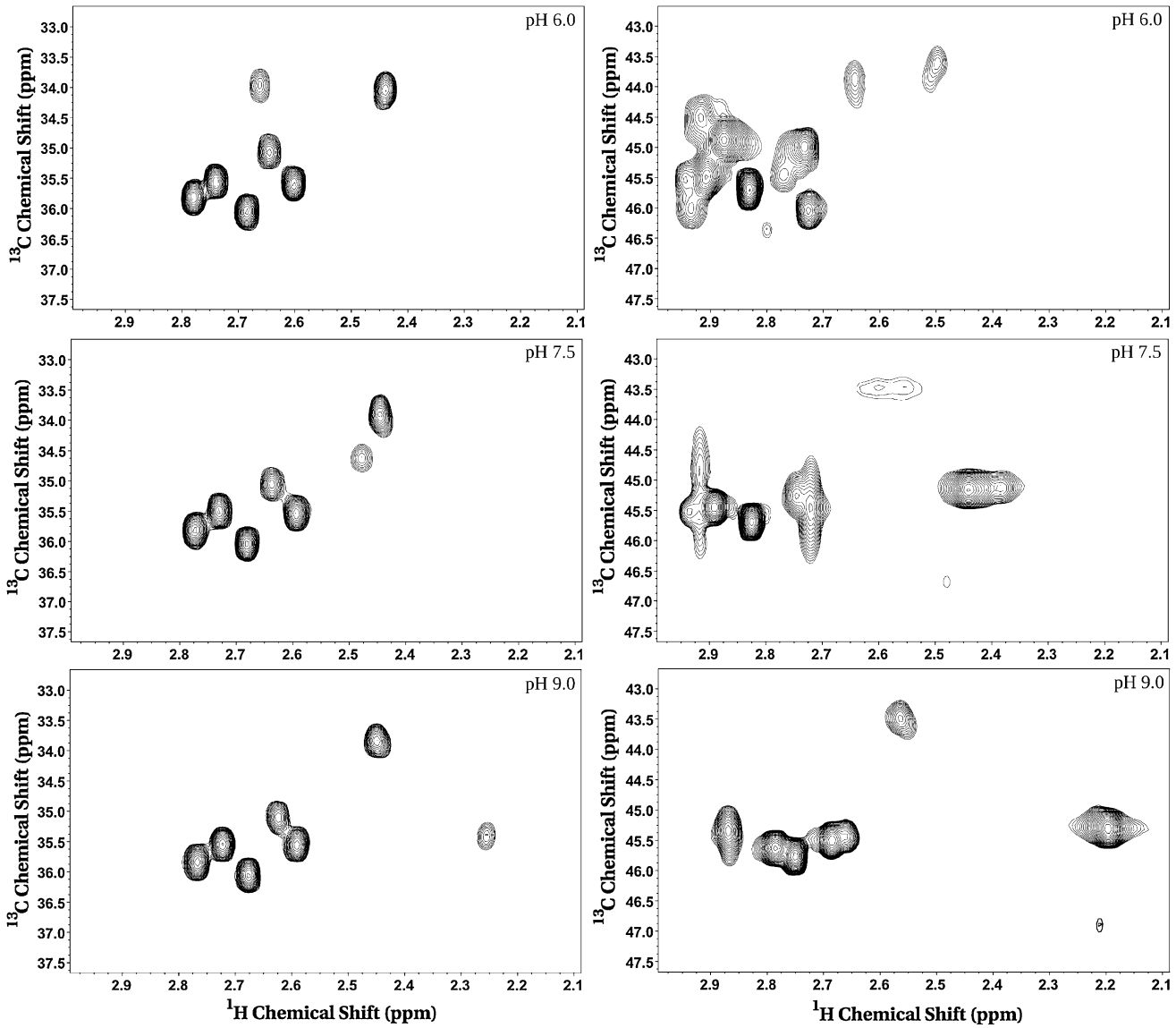

Fig. 3 (^{13}C,^{1}H) HSQCs of reductively methylated HEWL at 37 °C showing the mono- (*left*) and dimethyl regions (*right*) at pH 6.0, 7.5, and 9.0. Note that the K1(ε) dimethyl resonance, exists as two resonances at low pH (6.0) but coalesces to a single resonance with increasing pH

resonances near physiological pH. Moreover, given that mono- and dimethyl lysines exhibit comparable pKa's and deprotonation kinetics, we conclude that protonation kinetics alone cannot account for the line broadening. In other words, while deprotonation events are frequent, the populations are heavily weighted towards the protonated states. Therefore any broadening due to sampling of the deprotonated state would be negligible.

Goux et al. (1984) observed broadening of ^{13}C-dimethyl amino resonances in one-dimensional carbon spectra for reductively methylated concanavilin (ConA). They employed model compounds to show that the pH-dependence of line broadening could be adequately explained by base-catalyzed nitrogen inversions. At pH < 5.6, the ^{13}C methyl resonances of tetramethyllysine (TML) are represented by two distinct peaks, attributable to two distinct methyl conformers related through inversion at the nitrogen center. Above pH 5.6, the resonances coalesce, yet there still exists substantial broadening at near-physiological pH. At pH 10, the ^{13}C linewidths were sharp suggesting isotropic conditions or fast-regime exchange between methyl conformers. They explain that while they could not resolve ^{13}C re sonances for each of the ε-dimethyllysyl residues in ConA, they expect these to exhibit the same pH-dependent line broadening trend as for the N-terminal α-dimethylamino group whose inversion rate is on the order of 10^7 s^{-1} (Goux et al. 1984). With nitrogen inversions being as fast as they are, exchange between methyl conformers alone would be too rapid to give rise to line broadening. It is therefore apparent that transient protonation/deprotonation events serve to lock and unlock inversion at the

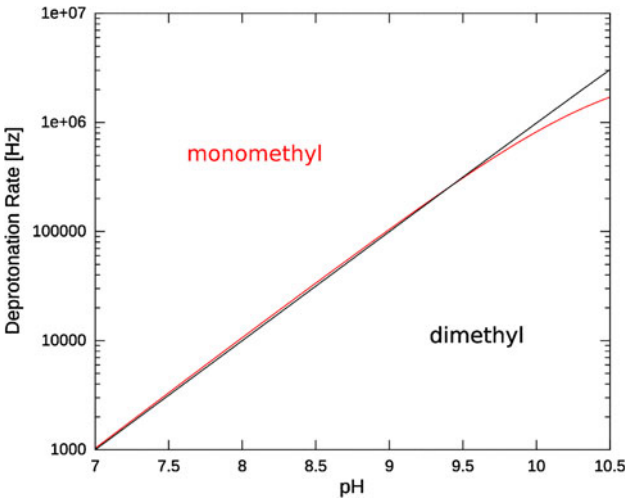

Fig. 4 Theoretical deprotonation rates for mono- and dimethyl lysines as a function of pH. Deprotonation rates are calculated per Lys species assuming a buffer concentration of 50 mM and buffer pKa of 7.2. Average monomethyl (11.9) and dimethyl (10.5) pKa were employed. We find that deprotonation rates are comparable among both species until pH > 10

nitrogen. The populations are heavily weighted toward the protonated species but at neutral pH, roughly once every millisecond, solvent-mediated deprotonation of the amine allows for sampling of the other methyl conformer(s). Protonation extends the lifetimes of the different methyl conformers, thereby putting the exchange between states in the intermediate NMR timescale regime. In addition to inversions, differences between each of the dimethylamino pKas, their solvent-exposure, and steric environments, serve to explain the observed variability in line broadening as a function of pH between the 8 resonances in the dimethyl spectra (Fig. 3).

Reaction schemes to favor monomethylated lysines

The scheme shown in Fig. 5 illustrates the key intermediates involved in the reductive methylation of lysine residues. The reaction mechanism can be described in terms of the formation of imine with subsequent conversion to the methylated species upon reduction (Means and Feeney 1968, 1995). The first pass through the cycle generates monomethyl while the second round produces dimethyl lysine.

Reductive methylation of lysine residues is strongly biased toward complete conversion to the dimethyl species. Following monoalkylation, the resulting secondary amine is considerably more nucleophilic than unlabeled, primary amine. Free formaldehyde, therefore, preferentially reacts with monomethyl lysines over unlabeled lysines, favoring dimethylation. Consequently, it becomes difficult to maintain a large population of monomethyl lysines while ensuring labeling at every reactive site. Our approach to monoalkylation is to employ a pH at which imine formation is most favorable and a low ratio of formaldehyde to amine, followed by reduction with a large excess of cyanoborohydride. Experimentally, methylation at pH ~6 enhances mono to dimethyl ratios substantially over reactions at higher pH (Fig. 6). When labeling at pH 7.0–7.5 with a stoichiometry of 1:1 to 1.5:1 (formaldehyde:amine), the mono to dimethyl peak integrals are roughly 1:1 in a given sample; with effective peak intensities of 1:2 (mono:dimethyl) given the twofold greater number of dimethyl protons. At higher formaldehyde:amine stoichiometry, as per the dimethylation

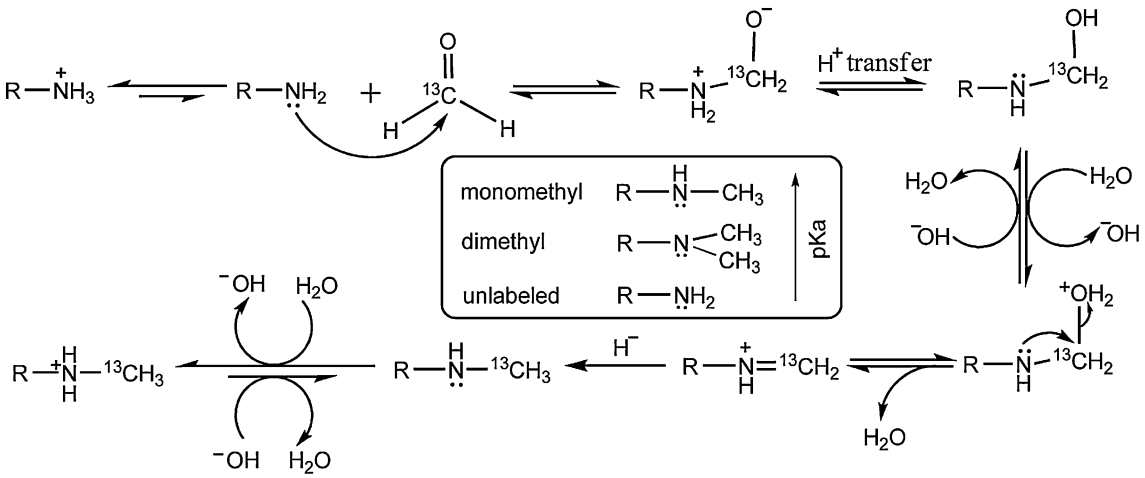

Fig. 5 Reaction scheme for 13C-reductive methylation of lysines. H- refers to any hydride donor suitable for reducing the iminium ion to the corresponding secondary (monomethyl) or tertiary amine (dimethyl); a variety of reducing agents can be employed. R represents the rest of the protein chain. Note that while the amine nucleophilicity does not necessarily correlate with pKa, the monomethyl species is considerably more nucleophilic than the primary, unlabeled amine

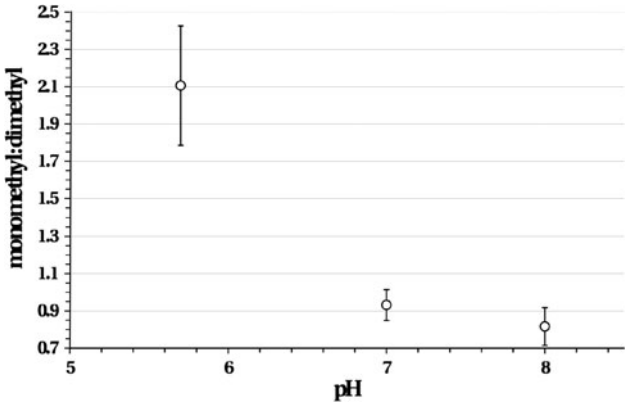

Fig. 6 Monomethyl to dimethyl ratios for HEWL labeled at pH 5.7, 7.0, and 8.0. Resonances were integrated using NMRViewJ (Johnson and Blevins 1994) and ratios obtained by comparing monomethyls to their corresponding dimethyl peak according to assignments in Fig. 1. The *data* represents the average of these ratios. *Vertical bars* represent standard error of the mono:di ratios for all lysine residues

protocol, mono to dimethyl ratios are substantially reduced. For the monoalkylation scheme we describe, roughly half of the signal arises from monomethyl derivatives (3 proton each) and half from dimethyl lysines (6 proton each); giving roughly a 2:1 ratio of mono:dimethyl.

The effect of aldehyde stoichiometry on the mono:-dimethyl ratios can be seen in Figures S8, S9, and S10, for methylation of HEWL (S8) and BSA (S9/S10) (Online Resource 8-10), respectively. The labeling efficiency at a given lysine is dependent on pKa and as would be expected, the K1α of HEWL most readily incorporates ^{13}C, often to the point of complete dimethylation even when the monoalkylation scheme is employed.

The reducing agent of choice also influences the dimethyl to monomethyl ratio. While use of large molar excess of cyanoborohydride does improve monomethyl yield, it also enhances byproduct formation; we observe direct reduction of formaldehyde to methanol. Overall, monomethylation is less efficient than dimethylation but gives substantial improvements in spectral quality.

^{13}C-deuteromethylation of protein amino groups

While ^{13}C-monomethylation of lysine residues provides improved resolution at physiological pH over that of dimethyl lysines, our experience has shown that dimethyl lysines serve as useful reporters of protein conformational exchange and dynamics. We sought to further improve the tags by incorporating a deuterium nucleus in the methyl groups. $^{13}CH_2D$ deuteromethyl probes provide a means to observe methyl lysines in macromolecular systems where natural abundance ^{13}C signal may otherwise obscure detection. Deuteromethylation can be achieved by modification of the standard reductive alkylation protocol

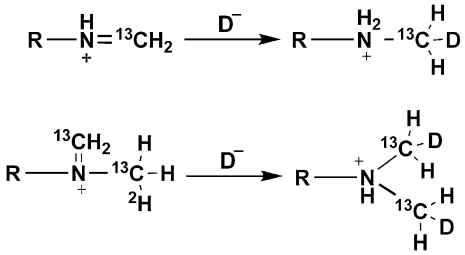

Fig. 7 Reduction of imine with deuteride generates mono- and di-deuteromethyl labels

whereby pre-formed imine is reduced with deuteride rather than hydride (Fig. 7).

Alternatively $^{13}CHD_2$-labeling is afforded by using ^{13}C,d_2 formaldehyde and a suitable mild reducing agent such as sodium cyanoborohydride or dimethylamine borane, whereas reduction with sodium cyanoborodeuteride would generate the corresponding $^{13}CD_3$ tag (Fig. 8). A CD_3 dimethyl lysine tag should prove equally effective in mass spectrometry applications, where isotope coded affinity tagging (ICAT) is utilized (Goshe and Smith 2003). In this case, tagging by a protonated or perdeuterated dimethyl tag would result in a difference of mass to charge ratios (Δm/z) of 6. Recently, Kashai et al. (2011) outlined an approach in which multiple ligand-specific conformations of the β_2-adrenergic receptor could be identified by monitoring the extent of reactivity of specific lysine residues on the membrane protein through ICAT labeling by a protonated and deuterated version of succinic anhydride. Reductive alkylation has the advantage that charge is preserved by the reaction, while the dimethyl tag is minimally perturbing, in which case the delicate ensemble of membrane protein conformers should be faithfully reproduced.

The versatility of reductive alkylation lies in the diversity of aldehydes, ketones and reducing agents that can be employed. For example, primary and secondary amines can also be tagged with ^{3}H via tritiated cyanoborohydride, thereby generating radionuclide labeled methyl groups. Figure 9 presents an overlay of spectra obtained from a standard ^{13}C,^{1}H HSQC versus that from a deuterium-edited version for $^{13}CH_2D$-deuteromethyl labeled apo-calmodulin in polyvinyl pyrrolidone (PVP-40). Deuterium editing was achieved by making use of the C-^{2}H coupling as described by Muhandiram et al. (1995). The PVP concentration was

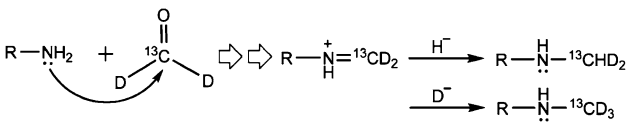

Fig. 8 $^{13}CH_2D$ and $^{13}CD_3$ deuteromethyl labeling is preceded by imine formation with ^{13}C,d_2 formaldehyde. Reduction with hydride generates a CHD2 label while deuteride affords CD_3 tags

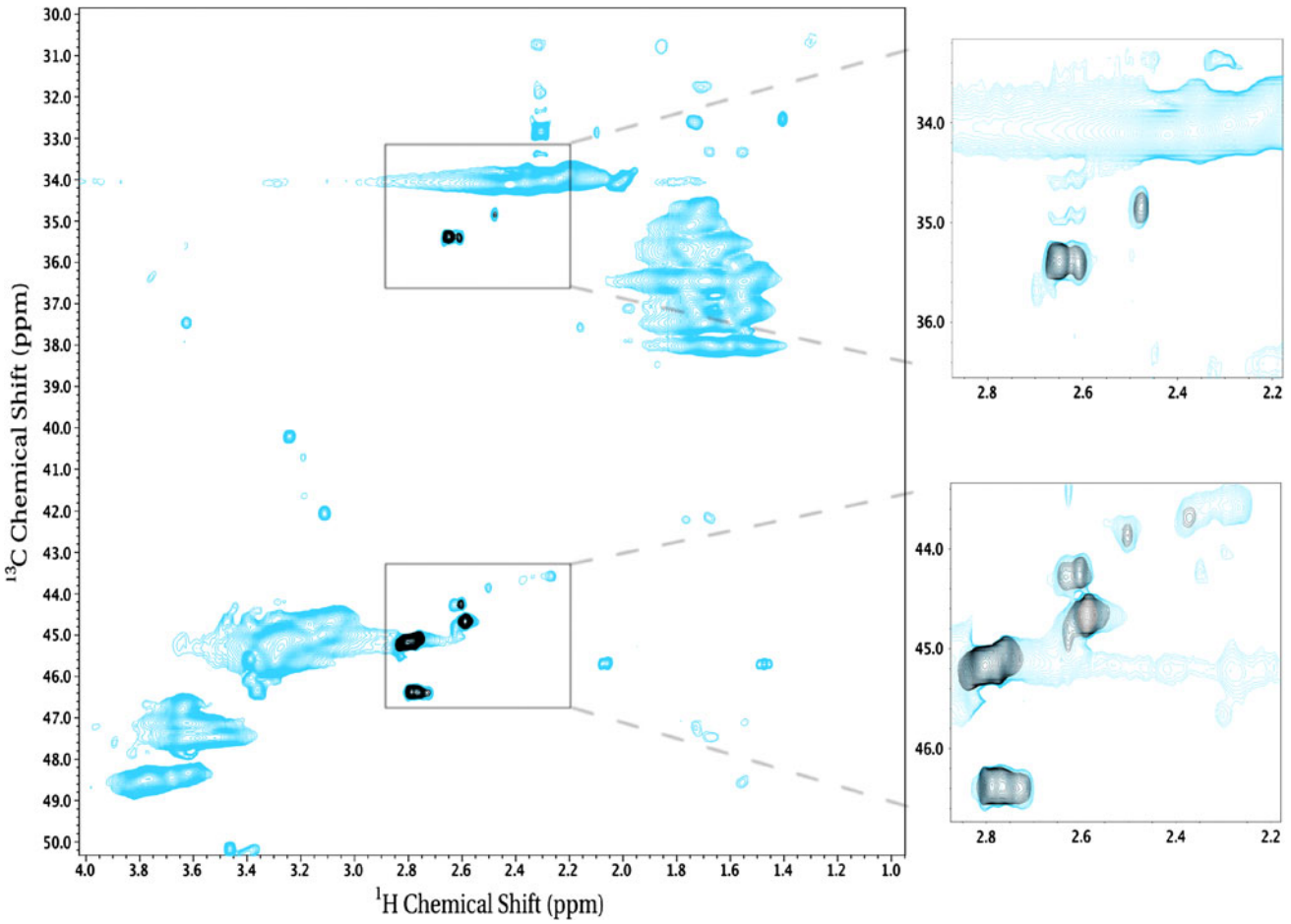

Fig. 9 Comparison of a standard ^{13}C,^{1}H HSQC spectrum of $^{13}CH_2D$-deuteromethylated apo-calmodulin in 300 mg/mL polyvinylpyrrolidone (*cyan*) versus a ^{2}H-edited ^{13}C,^{1}H HSQC spectrum (*black*) acquired on the same sample. Deuterium filtering eliminates all background PVP polymer signal and allows for unambiguous identification of the signal arising from the ^{13}C-deuteromethyl probes. Expansions of mono and dimethyl lysine regions show minor methyl lysine species that are well above baseline for the ^{2}H-edited spectrum but within the noise regime for the unedited HSQC

such that the macromolecular crowding approximates that which is found within cells (300 mg/mL).

The deuteromethyl tag and corresponding ^{2}H filtering schemes also provide improved water suppression while allowing for more detailed dynamics studies, as discussed by others, particularly with regard to CHD_2 probes (Tugarinov et al. 2005). Note from Fig. 9 that resonances for the methyllysines seem to be nearly as well resolved under conditions of high viscosity as they are under standard solution conditions.

Conclusions

Reductive methylation of primary amino groups with ^{13}C-formaldehyde is an effective means to generate site-specific methyl reporters on proteins. Aminomethyl probes can be used for the study of protein folding, dynamics, ligand binding and enzyme kinetics by NMR. However, without significantly raising the pH, it is found that dimethyl lysine resonances of HEW lysozyme exhibit pronounced line broadening, due to millisecond timescale reorientations at physiological temperatures. These intermediate timescale reorientations can be attributed primarily to salt-link breakage and reformation, and deprotonation-mediated nitrogen inversions. Other possible contributing factors to methylysine resonance broadening include: intra- and intermolecular hydrogen bonds, sterics, and hydrophobic interactions.

Significant line narrowing of dimethyllysines is often observed at a pH of 8.5 or greater where deprotonation rates, and consequently, nitrogen inversions, are fast on an NMR timescale. This rapid side chain reorientation is also found to be the case, at any pH, for monomethyl lysines since the single methyl is less restrictive toward bond rotations or inversions. More rapid interconversion between local anisotropic environments results in sharper line widths. Accordingly, certain non-salt-linked dimethyllysine resonances are sharper than others, presumably because of differences in their pKas and local environment.

While typical schemes for reductive alkylation of proteins with ^{13}C-formaldehyde generate dimethyl-lysines, monomethylation can be favored by suitable stoichiometry and pH. Use of roughly 1–1.2 equivalents of aqueous formaldehyde relative to primary amine is crucial to prevent over-modification of sites while at twofold to threefold equivalents or higher, the reaction is heavily weighted toward dimethylation. Nearly selective monomethylation can be achieved under conditions that favour imine formation and minimal labeling. While higher monomethyl yields are readily achieved using the protocol described above, it becomes inevitable that some sites will be dimethylated (especially for the α-NH_2 which has a significantly lower pKa than the ε-NH_2 species). We have demonstrated a modification to conventional reductive alkylation and presented applications of ^{13}C-deuteromethylation to studies of proteins by NMR. Benefits of ^{13}C-deuteromethylation in combination with ^{2}H-editing of ^{13}C,^{1}H HSQCs include: (1) enhanced spectral quality via elimination of any and all background ^{13}C-coupled proton signal, (2) unambiguous identification of signal arising from the aminomethyl probes, (3) substantial minimization of water signal breakthrough, and (4) the possibility of studying side chain dynamics via deuterium relaxation and ^{2}H CPMG dispersion measurements. Finally, ^{13}C-methyl probes on lysines are less susceptible to line-broadening associated with slower molecular tumbling given the high degree of rotational correlation of the sidechain. Methyl lysines should therefore serve useful as probes for *in cell* spectroscopy studies of proteins, where the protein of interest is either injected or transported into the cell. We emphasize that reductive ^{13}C methylation is a robust, mild labeling protocol that can be applied under non-denaturing conditions to virtually any protein. Reaction conditions can be tailored to preferentially monoalkylate while isotopically labeled formaldehyde and or reducing agents provide further utility to the tags, in the case of $^{13}CH_2D$, $^{13}CHD_2$, and $^{13}CD_3$ deuteromethyls.

Acknowledgments RSP acknowledges NSERC (Grant number 261980) for a research discovery award. MPB acknowledges support from the Stanford Medical Scientist Training Program.

References

Abe Y, Ueda T, Imoto T (1997) An improved method for preparing lysozyme with chemically C-13-enriched methionine residues using 2-aminothiophenol as a reagent of thiolysis. J Biochem 122:1153–1159

Abraham SJ, Kobayashi T, Solaro RJ, Gaponenko V (2009) Differences in lysine pKa values may be used to improve NMR signal dispersion in reductively methylated proteins. J Biomol NMR 43:239–246

Antos J, Francis M (2006) Transition metal catalyzed methods for site-selective protein modification. Curr Opin Chem Biol 10: 253–262

Bokoch MP, Zou Y, Rasmussen SGF, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJ, Thian FS, Kobilka TS, Puglisi JD, Weis WI, Pardo L, Prosser RS, Mueller L, Kobilka BK (2010) Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. Nature 463:108–112

Borisenko V, Sansom M, Woolley GA (2000) Protonation of lysine residues inverts cation/anion selectivity in a model channel. Biophys J 78:1335–1348

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRpipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Gerken T, Jentoft J, Jentoft N, Dearborn D (1982) Intramolecular interactions of amino groups in 13C reductively methylated hen egg-white lysozyme. J Biol Chem 257:2894–2900

Goshe MB, Smith RD (2003) Stable isotope-coded proteomic mass spectrometry. Curr Opin Biotechnol 14:101–109

Goux WJ, Teherani J, Sherry AD (1984) Amine inversion in proteins—a C-13-NMR study of proton-exchange and nitrogen inversion rates in n-epsilon, n-epsilon, n-alpha, n-alpha-C-13 tetramethyllysine, n-epsilon, n-epsilon, n-alpha, n-alpha C-13 tetramethyllysine methyl-ester, and reductively methylated concanavalin-A. Biophys Chem 19:363–373

Ikura M, Kay LE, Bax A (1990) A novel-approach for sequential assignment of H-1, C-13, and N-15 spectra of larger proteins—heteronuclear triple-resonance 3-dimensional NMR-spectroscopy—application to calmodulin. Biochemistry 29:4659–4667

Imoto T, Johnson LN, North ACT, Phillips DC, Rupley JA (1972) Lysozyme in enzymes. Academic Press, New York

Johnson BA, Blevins RA (1994) NMR view: a computer-program for the visualization and analysis of NMR data. J Biomol NMR 4:603–614

Joshi N, Whitaker L, Francis M (2004) A three-component Mannich-type reaction for selective tyrosine bioconjugation. J Am Chem Soc 126:15942–15943

Kahsai AW, Rajagopal S, Ahn S, Shukla AK, Sun J, Oas TG, Xiao K, Lefkowitz RJ (2011) Multiple ligand-specific conformations of the beta(2)-adrenergic receptor. Nat Chem Biol 7:692–700

Kalbitzer HR, Rohr G, Nowak E, Goody RS, Kuhn W, Zimmermann H (1992) A new high-sensitivity F-19 probe for labeling cysteine groups of proteins. NMR Biomed 5:347–350

Klein-Seetharaman J, Getmanova E, Loewen M, Reeves P, Khorana H (1999) NMR spectroscopy in studies of light-induced structural changes in mammalian rhodopsin: applicability of solution 19F NMR. Proc Natl Acad Sci USA 96:13744–13749

Korzhnev DM, Mittermaier AK, Kay LE (2005) Cross-correlated spin relaxation effects in methyl H-1 CPMG-based relaxation dispersion experiments: complications and a simple solution. J Biomol NMR 4:337–342

Kumar S, Nussinov R (1999) Salt bridge stability in monomeric proteins. J Mol Biol 293:1241–1255

Kumar S, Tsai CJ, Ma BY, Nussinov R (2000) Contribution of salt bridges toward protein thermostability. J Biomol Struct Dyn 17(suppl 1):79–85

Luchette P, Prosser R, Sanders C (2002) Oxygen as a paramagnetic probe of membrane protein structure by cysteine mutagenesis and F-19 NMR spectroscopy. J Am Chem Soc 124:1778–1781

Macnaughtan MA, Kane AM, Prestegard JH (2005) Mass spectrometry assisted assignment of NMR resonances in reductively C-13-methylated proteins. J Am Chem Soc 127(50):17626–17627

Means GE, Feeney RE (1968) Reductive alkylation of amino groups in proteins. Biochemistry 7:2192–2201

Means GE, Feeney RE (1995) Reductive alkylation of proteins. Anal Biochem 224:1–16

Moore G, Cox M, Crowe D, Osborne M, Rosell F, Bujons J, Barker P, Mauk M, Mauk A (1998) N epsilon, N epsilon-dimethyl-lysine cytochrome c as an NMR probe for lysine involvement in protein–protein complex formation. Biochem J 332:439–449

Moult J, Yonath A, Traub W, Smilansky A, Podjarny A, Rabinovich D, Saya A (1976) Structure of triclinic lysozyme at 2.5 a resolution. J Mol Biol 100:179–195

Muhandiram D, Yamazaki T, Sykes B, Kay L (1995) Measurement of H-2 T-1 and T-1p relaxation-times in uniformly C-13-labeled and fractionally H-2-labeled proteins in solution. J Am Chem Soc 117:11536–11544

Ojida A, Tsutsumi H, Kasagi N, Hamachi I (2005) Suzuki coupling for protein modification. Tetrahedron Lett 46:3301–3305

Oxenoid K, Sonnichsen FD, Sanders CR (2002) Topology and secondary structure of the N-terminal domain of diacylglycerol kinase. Biochemistry 41:12876–12882

Palmer AG, Kroenke CD, Loria JP (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods Enzymol 339:204–238

Rayment I (1997) Reductive alkylation of lysine residues to alter crystallization properties of proteins. Methods Enzymol 276: 171–179

Rypniewski W, Holden H, Rayment I (1993) Structural consequences of reductive methylation of lysine residues in hen egg white lysozyme: an X-ray analysis at 1.8-Anstrom resolution. Biochemistry 32:9851–9858

Sarakatsannis J, Duan Y (2005) Statistical characterization of salt bridges in proteins. Proteins Struct Funct Bioinf 60:732–739

Sprangers R, Velyvis A, Kay L (2007) Solution NMR of supramolecular complexes: providing new insights into function. Nat Methods 4:697–703

Tieleman DP, Borisenko V, Sansom MSP, Woolley GA (2003) Understanding pH-dependent selectivity of alamethicin K18 channels by computer simulation. Biophys J 84:1464–1469

Tsai CS, Tsai YH, Lauzon G, Cheng ST (1974) Structure and activity of methylated horse liver alcohol-dehydrogenase. Biochemistry 13:440–443

Tugarinov V, Kay LE (2006) A 2H NMR relaxation experiment for the measurement of the time scale of methyl side-chain dynamics in large proteins. J Am Chem Soc 128:12484–12489

Tugarinov V, Ollerenshaw JE, Kay LE (2005) Probing side-chain dynamics in high molecular weight proteins by deuterium NMR spin relaxation: an application to an 82-kDa enzyme. J Am Chem Soc 127:8214–8225

White HD, Rayment I (1993) Kinetic characterization of reductively methylated myosin subfragment 1. Biochemistry 32:9859–9865

Yokoyama S (2003) Protein expression systems for structural genomics and proteomics. Curr Opin Chem Biol 7:39–43

Zhang M, Vogel H (1993) Determination of the side chain pKa values of the lysine residues in calmodulin. J Biol Chem 268:22420

Zhang MJ, Thulin E, Vogel HJ (1994) Reductive methylation and pKa determination of the lysine side-chains in calbindin D-9K. J Protein Chem 13:527–535