NMR Analysis of Methyl Groups at 100−500 kDa: Model Systems and Arp2/3 Complex†

Mara Kreishman-Deitrick,‡ Coumaran Egile,§ David W. Hoyt,|| Joseph J. Ford,|| Rong Li,§ and Michael K. Rosen*,‡

Department of Biochemistry, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, Texas 75390, Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, and Pacific Northwest National Laboratory, Environmental Molecular Sciences Laboratory, 3335 Q Avenue, Richland, Washington 99352

Received April 4, 2003; Revised Manuscript Received May 14, 2003

ABSTRACT: Large macromolecular machines are among the most important and challenging targets for structural and mechanistic analyses. Consequently, there is great interest in development of NMR methods for the study of multicomponent systems in the 50−500 kDa range. Biochemical methods also must be developed in concert to produce such systems in selectively labeled form. Here, we present $^{1}H/^{13}C$-HSQC spectra of protonated methyl groups in a model system that mimics molecular weights up to ∼560 kDa. Signals from side chain methyl groups of Ile, Leu, and Val residues are clearly detectable at correlation times up to ∼330 ns. We have also developed a biochemical procedure to produce the 240 kDa, heteroheptameric Arp2/3 actin nucleation complex selectively labeled at one subunit and obtained $^{1}H/^{13}C$-HSQC spectra of this assembly. Sensitivity in spectra of both the Arp2/3 complex and the model system indicate that methyl groups will be useful sources of information in nonsymmetric systems with molecular weights greater than 600 kDa at concentrations less than 100 μM. Methyl analyses will complement TROSY and CRINEPT analyses of amides in NMR studies of structure and molecular interactions of extremely large macromolecules and assemblies.

Biomolecular solution NMR is an extremely powerful tool with which to study the structure, conformational dynamics, and ligand interactions of macromolecules and assemblies. A significant obstacle facing NMR spectroscopy of large systems is the rapid deterioration of NMR signals because of fast transverse relaxation. Because the rate of transverse relaxation is proportional to the rotational correlation time (τ_C) of a molecule, signal loss resulting from it becomes severe as molecular weight increases. The recent development of transverse relaxation-optimized spectroscopy (TROSY),[1] in combination with cross relaxation-induced polarization transfer (CRIPT) and cross relaxation-enhanced polarization transfer (CRINEPT) techniques, has overcome many of the deleterious effects of transverse relaxation of amide nuclei during polarization transfer, evolution, and acquisition periods of NMR pulse sequences (*1*, *2*). These pulse sequences have been used to determine global folds, measure dipolar couplings, and probe ligand interactions in systems with molecular weights in the range of 40−60 kDa (*3*−*6*). NMR backbone assignments have been obtained using TROSY-based triple resonance experiments for proteins ranging up to 110 kDa (*4*, *6*, *7*). CRIPT-TROSY sequences have most recently been used to obtain two-dimensional NMR data on the 900 kDa, 14-fold symmetric GroEL−GroES chaperone complex, which marks a significant advancement in the application of solution NMR methods to extremely large macromolecules (*8*).

The above work has focused on NMR signals from protein amide groups. Side chain methyl groups offer a useful alternative spectroscopic probe in large systems. Methyl groups have several advantages over amides in terms of their utility in NMR spectroscopy. Methyl protons are not exchangeable with solvent, whereas amide proton line widths are inherently broadened by solvent exchange. ^{1}H-^{13}C correlation spectra are generally disperse and well-resolved, and methyl ^{13}C and ^{1}H lines are inherently narrower as a result of rapid rotation of the three protons about the methyl symmetry axis (*9*−*11*). The presence of three protons per methyl group provides higher sensitivity and efficient longitudinal relaxation through intra-methyl dipolar interactions, resulting in relatively rapid T1 relaxation, even in systems highly deuterated at nonmethyl positions (*10*, *12*). It has been shown that for $^{1}H,^{13}C$(methyl)/U-$^{15}N,^{2}H$-labeled samples of maltose binding protein (42 kDa) and *Staphylo-*

† This work was supported by NIH Grant R01 GM56322 (M.K.R.), the Human Frontier Science Program (C.E.), and NIH Grant R01 GM057063 (R.L.).

* To whom correspondence should be addressed. Phone: (214) 648-4703. Fax: (214) 648-8856. E-mail: mrosen@biochem.swmed.edu.

‡ University of Texas Southwestern Medical Center at Dallas.

§ Harvard Medical School.

|| Pacific Northwest National Laboratory.

[1] Abbreviations: Arp2/3 complex, actin-related protein 2/3 complex; CRINEPT, cross relaxation-enhanced polarization transfer; CRIPT, cross relaxation-induced polarization transfer; GDID59, residues 60−204 of bovine Rho guanine nucleotide dissociation inhibitor; GST, glutathione *S*-transferase; HSQC, heteronuclear single-quantum coherence; IgG, immunoglobulin G; INEPT, insensitive nuclei enhanced by polarization transfer; NMR, nuclear magnetic resonance; N-WASP, neuronal Wiskott−Aldrich Syndrome protein; S/N, signal-to-noise ratio; TROSY, transverse relaxation-optimized spectroscopy; VCA peptide, verprolin homology, central, acidic region peptide; WASP, Wiskott−Aldrich Syndrome protein.

10.1021/bi034536j CCC: $25.00
Published on Web 06/25/2003

coccus aureus 7,8-dihydroneopterin aldolase (DHNA, 110 kDa), a simple $^{1}H/^{13}C$-HSQC spectrum of the methyl resonances is 3- and ~10-fold more sensitive, respectively, than a $^{1}H/^{15}N$-TROSY spectrum of the amides in the same sample (*13*). Methyl groups are also enriched in protein interiors and at protein−protein interfaces (*14*). As with amide moieties, perturbation of methyl chemical shifts can be used to qualitatively map ligand interaction surfaces and report on conformational changes, information useful in structure determination and mechanistic studies (*13*, *15*).

Large, multicomponent assemblies represent some of the most important targets for structural and biophysical investigations. NMR analysis of such systems requires both biochemical methods to prepare them in labeled form and NMR methods to address their high molecular weight. The systems used to develop TROSY/CRIPT-based sequences (DHNA and GroEL/ES), while large (110 and 900 kDa, respectively), are both symmetrical homooligomers. This results in spectra limited to a single set of degenerate signals originating from each symmetry-related monomer, with the effective monomer concentration being manyfold higher than that of the assembly. The ability to produce these systems in *Escherichia coli* through expression of a single chain also eliminates the need for selective expression and labeling of separate components. In these respects, DHNA and GroEL/ES, while useful models, are not representative of the majority of macromolecular assemblies, which lack high internal symmetry. The Arp2/3 actin nucleation complex is a 240 kDa asymmetric assembly that is more representative of the large class of multimeric macromolecular machines. Arp2/3 complex is a heteroheptamer consisting of subunits Arp2, Arp3, p40, p35, p21, p20, and p16 (mammalian nomenclature). Only one copy of each subunit is present in each assembly, resulting in a monomer concentration limited by the solubility of the complex (~100 μM). The intact assembly cannot be produced recombinantly in bacteria. Moreover, since the signals from each subunit are distinct, Arp2/3 complex will require reconstitution from separately expressed and labeled parts to achieve a labeling pattern that will limit the number of NMR-visible nuclei to those within the subunit(s) of interest. This is likely to be the case in the majority of multimeric biological complexes. Finally, recent work has revealed that regulation of the actin nucleating activity of Arp2/3 complex by the Wiskott−Aldrich Syndrome protein (WASP) family proteins and filamentous actin is likely to involve a significant quaternary rearrangement in the complex (*16*, *17*). The crystal structure of the inactive form of the assembly is known (*16*), although the nature of the activating conformational change and the binding modes of WASP proteins and actin filaments are not currently understood.

In this study, we explore the limits of utility of protonated side chain methyl groups as NMR probes in very large systems, as well as the biochemical methods necessary to reconstitute such systems in selectively labeled form. Using a model system with estimated rotational correlation times up to ~330 ns (corresponding to a molecular weight of ~560 kDa), we find that signals from well-resolved methyl resonances can be detected in a 4.5-h $^{1}H/^{13}C$-HSQC experiment on a 0.8 mM model sample even at $\tau_C \approx 330$ ns. Intense signals can also easily be detected in a 96-h spectrum of ~100 μM Arp2/3 complex selectively labeled at the p21 subunit. These results indicate that protonated methyl groups will be powerful tools for the study of large, reconstituted macromolecular systems by NMR.

EXPERIMENTAL PROCEDURES

Yeast Strain. Saccharomyces cerevisiae strain CEY64 (*MATa ura3-52,his3-200,leu2-3,lys2-801,ARP3::(ARP3-ProA2,URA3),arc18::HIS3*) was constructed as follows. The *arc18* gene was deleted by replacement with a *HIS3* gene. Affinity tagging of Arp3p with two IgG binding domains of *Staphylococcal* Protein A was accomplished by a one-step insertion of (*ProA2,URA3*)K7 at the 3′ untranslated region of the *ARP3* open reading frame.

Protein Expression and Purification. $^{13}C/^{15}N/^{2}H/^{1}H$(methyl) GDIΔ59 and p21 were expressed in BL21(DE3) *E. coli* cells as described (*9*). $^{13}C/^{15}N/^{2}H/^{1}H$(methyl) and $^{2}H/^{15}N$ GDIΔ59 were purified as described previously (*18*) and concentrated to 1.33 mM in NMR buffer (20 mM phosphate, pH 6.8, 50 mM NaCl, 1 mM EDTA, 2 mM DTT, 10% D_2O). ^{2}H-labeled glycerol was added to 46% (wt/v), resulting in a final protein concentration of 0.8 mM. $^{13}C/^{15}N/^{2}H/^{1}H$(methyl) p21 was purified using anion exchange (DEAE Sepharose), cation exchange (Mono-S), and size exclusion (Superdex 75) chromatographies. Selectively labeled Arp2/3 complex was purified as follows. A total of 1 kg of CEY64 yeast cells was resuspended in 1 L of ice cold lysis buffer (50 mM HEPES, pH 7.5, 100 mM KCl, 4 mM EGTA, 10 mM $MgCl_2$, 1 mM EDTA, 0.1% Brij58, 10% glycerol, 0.2 mM ATP, 1 mM PMSF, 6 mM DTT, protease inhibitor cocktail) and lysed using a DynoMill (Glen Mills, Inc.). Functional yields depended critically on sufficient heat exchange to maintain the sample at 0−4 °C throughout lysis. The lysate was cleared by centrifugation at 16 000*g* for 30 min, then at 138 000*g* for 1 h. A 10-fold molar excess (based on estimates of expected Arp2/3 complex yield) of purified p21 was added to the lysate, followed by incubation for 2 h at 4 °C. The lysate was incubated in a batch with 15 mL of IgG Sepharose for 1 h at 4 °C, followed by washing of the resin with lysis buffer. Protocols based on addition of p21 after adsorption of Arp2/3 complex to IgG Sepharose produced significantly lower yields. Arp2/3 complex was cleaved off the beads by incubation with 3000 units of Tev protease for 5 h at room temperature. The conductivity of the Tev reaction was lowered by dilution in SA buffer (10 mM PIPES, pH 6.8, 1 mM EGTA, 1 mM $MgCl_2$, 10% glycerol), and the cleaved complex was then applied to a cation exchange column (Mono-S) followed by Superdex 200 in NMR buffer (25 mM phosphate, pH 7.0, 50 mM NaCl, 1 mM EGTA, 1 mM $MgCl_2$, 1 mM DTT, 10% glycerol). The purified complex was concentrated and exchanged into NMR buffer containing 10% ^{2}H-labeled glycerol (glycerol-d_8, Cambridge Isotope Laboratories), resulting in a final concentration of 118 μM in a volume of 250 μL.

Proteins for pyrene actin polymerization assays were expressed and purified as follows. N-WASP VCA was expressed as a GST fusion protein in BL21(DE3) *E. coli* cells in LB medium and purified using anion exchange chromatographies (DEAE Sepharose and Mono-Q), followed by thrombin protease cleavage of the GST tag. The cleaved material was further purified using anion exchange (Mono-Q) and size exclusion (Superdex 75). The second Mono-Q

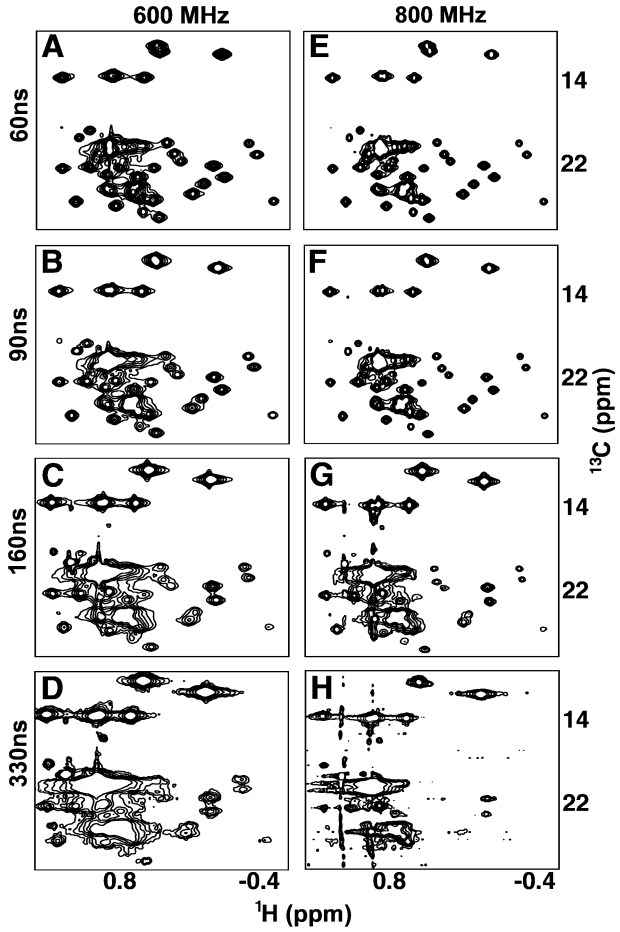

FIGURE 1: ^{1}H/^{13}C-HSQC spectra of GDIΔ59 in 46% (wt/v) glycerol. Spectra were recorded at 25 °C (A and E, $\tau_C \approx 60$ ns), 15 °C (B and F, $\tau_C \approx 90$ ns), 0 °C (C and G, $\tau_C \approx 160$ ns), and −10 °C (D and H, $\tau_C \approx 330$ ns) at either 600 MHz (A−D) or 800 MHz (E−H). Total acquisition time was 4.5 h per spectrum.

column also served to purify away any VCA that had degraded during initial purification and cleavage. $(His)_{10}$-p21 was expressed in BL21(DE3) *E. coli* cells in LB medium and purified using Ni-NTA agarose (Qiagen) followed by cation exchange (Mono-S) and size exclusion (Superdex 75). Bovine and yeast Arp2/3 complexes (wild-type and Δyp21), yeast GST-VCA, and rabbit skeletal muscle actin were purified as described (*19*−*23*). Unless otherwise indicated, all chromatography resins were obtained from Amersham Biosciences.

Pyrene Actin Polymerization Assays. Actin polymerization was monitored in a pyrene actin fluorescence assay, in which the rate of actin polymerization is monitored as an increase in pyrene fluorescence over time (*24*). Yeast Arp2/3 complex assays (*20*) were performed in the following buffer: 50 mM HEPES, pH 7.5, 100 mM KCl, 3 mM $MgCl_2$, 1 mM EGTA. Bovine Arp2/3 complex assays were performed as described (*21*). Concentrations of various components used in the assays are listed in the legend to Figure 1.

NMR Spectroscopy. ^{1}H/^{13}C-HSQC experiments on Arp2/3 complex were carried out at 25 °C on Varian Inova 600 and 800 MHz spectrometers with four channels and pulsed field gradients, with sweep widths of 15 ppm (^{1}H) and 45.6 ppm (^{13}C), t_2 and t_1 acquisition times of 64 and 14 ms, respectively, an INEPT transfer delay of 1.8 ms, and a total recycle delay of 0.9 s (96 h acquisition time). INEPT and recycle delays were optimized empirically. ^{1}H/^{13}C-HSQC experiments on GDIΔ59 were carried out at temperatures ranging from 25 to −10 °C on Varian Inova 600 and 800 MHz spectrometers with sweep widths of 15 ppm (^{1}H) and 45.6 ppm (^{13}C), t_2 and t_1 acquisition times of 64 and 14 ms, respectively, and a total recycle delay of 1.0 s (4.5 h acquisition time). ^{1}H/^{15}N-TROSY and ^{1}H/^{15}N-CRINEPT-TROSY experiments were acquired at 10 °C on a Varian Inova 800 MHz spectrometer. All TROSY-based experiments were recorded with sweep widths of 12 ppm (^{1}H) and 28 ppm (^{15}N), t_2 and t_1 acquisition times of 64 and 56 ms, respectively, and a total recycle delay of 1.5 s (4.5 h acquisition time). All spectra were processed in NMRPipe using a $\cos^2$ window function. S/N and line width analyses were performed using NMRDraw. NMR assignments of GDIΔ59 were obtained previously (*18*).

Table 1: Solution Viscosities, Estimated Correlation Times, and Apparent Molecular Weights of 0.8 mM GDIΔ59 in 46% (wt/v) Glycerol

T (°C)	$\eta_{46\%}$ (cP)[a]	$\eta_{46\%} + \Delta 59$ (cP)[b]	estimated τ_C (ns)[c]	approximate MW (kDa)[d]
25	4.17	4.40	60	100
20	4.95	5.25	70	120
15	5.86	6.21	90	150
10	6.93	7.34	110	190
5	8.19	8.68	130	220
0	9.69	10.27	160	270
−5	15	16	270	460
−10	20	21	330	560

[a] Glycerol solution viscosities at 25−0 °C were calculated from values between 20 and 30 °C, assuming $\ln(\eta)$ is linear with temperature (*26*). Viscosities at −5 and −10 °C were interpolated from values at these temperatures for 40 and 50 wt %/v glycerol (*27*). [b] Total solution viscosities including 0.8 mM GDIΔ59 were calculated using the Einstein−Simha equation for the viscosity of suspensions (*41*). [c] Estimated τ_C values were calculated using Stokes' Law (*42*) based on a measured τ_C of 12.1 ns in aqueous buffer ($\eta_{0\%} = 0.893$). [d] Molecular weights were estimated from $\tau_C = 0.585$ (MW) (see Supporting Information Figure 1).

RESULTS

^{1}H/^{13}C-HSQC Spectra for τ_C Up to 330 ns. To assess the general applicability of methyl groups as NMR probes in high molecular weight systems, we developed a model system in which increasing molecular weight is simulated by modulating rotational correlation time (τ_C) with glycerol solution viscosity. As our model protein, we chose a 16 kDa C-terminal fragment of the protein RhoGDI (GDIΔ59) (*18*). In aqueous buffer (0% glycerol) at 25 °C, ^{15}N T_1 and T_2 relaxation times (*25*) indicate a correlation time for this protein of 12.1 ns (data not shown). In 46% (wt/v) glycerol, the correlation time of GDIΔ59 can be tuned between 60 and 330 ns (Table 1), approximating a molecular weight range of 100−560 kDa (Supporting Information Figure 1), by varying the temperature between 25 and −10 °C (*26*, *27*).

We acquired ^{1}H/^{13}C-HSQC spectra of ^{13}C,^{15}N,^{2}H,^{1}H-(methyl) GDIΔ59 in 46% (wt/v) glycerol at a range of temperatures at 800 and 600 MHz (Figure 1). The most striking feature of these spectra is that nearly all resonances can be observed even at the longest τ_C. Signals from Val and Leu side chains are all present at reasonable intensity, and well-resolved resonances can be readily identified. The

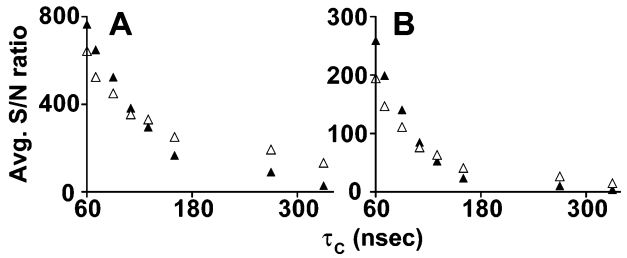

FIGURE 2: Analysis of S/N for increasing rotational correlation times. (A and B) Average S/N ratios at 800 MHz (▲) and 600 MHz (△) for five of seven Ile δ1 peaks (A) and nine selected Leu and Val methyl peaks (B) are plotted as a function of estimated τ_C (ns).

Ile δ_1 resonances remain very intense and clearly resolved even at −10 °C ($\tau_C \approx 330$ ns, Figure 1D,H), with the exception that line broadening causes the coalescence of the Ile 198 and Ile 114 signals into a single peak. The Ile δ1 signals are consistently more intense than those from the side chains of Leu and Val residues. The central region of the spectrum is very crowded, and line broadening makes it impossible to resolve individual lines, but the well-resolved resonances from the side chains of Leu 75, Leu 88, Leu 90, Leu 94, Leu 190, Leu 104, and Val 146 are observable at −10 °C. Notably, signals from methyl-containing side chains both in the core and on the surface of GDIΔ59 are observable in these spectra (*18*).

Average signal-to-noise (S/N) ratios for isolated Ile, Leu, and Val residues are plotted in Figure 2 A,B. As expected, sensitivity decreases significantly as the apparent molecular weight increases. The S/N reaches ∼ 4−15 for Leu and Val residues at $\tau_C \approx 330$ ns, while S/N for Ile residues remains in the range of ∼30−130, even at the longest observed correlation time. The stronger signal for Ile is most likely due to the absence of dipolar interactions between protons on adjacent diastereotopic methyl groups that occur in Val and Leu. Additional benefits may arise from increased mobility of the Ile methyl group relative to those of Val and Leu (*28*). At both 600 and 800 MHz, ^{1}H and ^{13}C line widths of Ile remain narrow. For Leu and Val, ^{1}H line widths broaden significantly with increasing τ_C (not shown). ^{13}C line widths could not be measured accurately with the resolution at which these spectra were acquired. Interestingly, sensitivity for all peaks is higher at 800 MHz than at 600 MHz for low τ_C values, but sensitivity becomes better at 600 MHz for τ_C greater than ∼130 ns. This effect can likely be attributed to the fact that at low τ_C sensitivity is inherently higher at 800 MHz, but as τ_C increases, ^{1}H T_1 relaxation times get longer, and this phenomenon is more severe at 800 MHz. At long τ_C then, the gain in inherent sensitivity at 800 MHz is outweighed by shorter T_1 relaxation times at 600 MHz.

^{1}H/^{15}N-TROSY and ^{1}H/^{15}N-CRINEPT-TROSY Spectra of a Methyl Protonated Sample. Having established that methyl signals can be readily detected in ^{13}C,^{15}N,^{2}H,^{1}H(methyl) GDIΔ59 even at very long τ_C values, we wished to investigate whether we could simultaneously study backbone amides in the same sample. We compared ^{1}H/^{15}N-TROSY and ^{1}H/^{15}N-CRINEPT-TROSY spectra acquired with ^{13}C,^{15}N,^{2}H,^{1}H(methyl) GDIΔ59 to spectra acquired with uniformly ^{15}N/^{2}H-labeled GDIΔ59 at 10 °C ($\tau_C \approx 110$ ns). Figure 3 shows that the sensitivity of both ^{1}H/^{15}N correlation

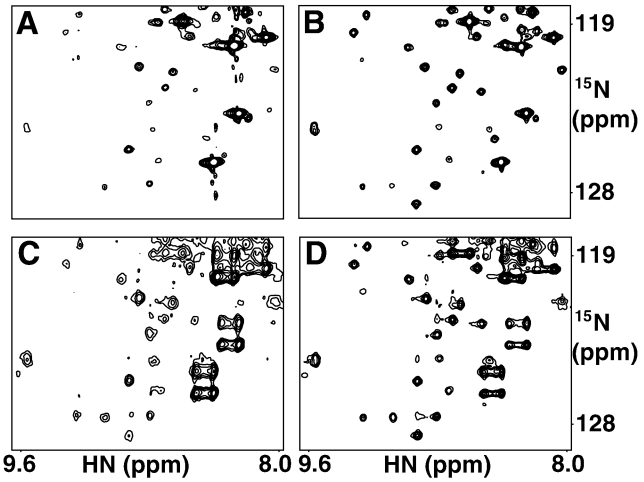

FIGURE 3: TROSY-based spectra of GDIΔ59 in 40% glycerol. Regions from ^{1}H/^{15}N-TROSY (A and B) and ^{1}H/^{15}N-CRINEPT-TROSY (C and D) spectra of ^{13}C/^{15}N/^{2}H/^{1}H(methyl) GDIΔ59 (A and C) and ^{15}N/^{2}H GDIΔ59 (B and E) are shown. Spectra were recorded in 4.5 h at 10 °C.

spectra is significantly degraded by the introduction of methyl protons into the uniformly deuterated background. Degradation is not uniform across the spectra. As expected, intensity is most severely decreased for amides proximal to methyl-containing sites. These findings contrast with previous results obtained with the 42 kDa maltose binding protein, where similar sensitivity was achieved in ^{1}H/^{15}N-TROSY spectra for fully deuterated and methyl protonated samples (*13*). Dipolar interactions between the backbone amide protons and the methyl protons have a much more severe effect on TROSY efficiency in very large systems than those of more moderate molecular weight. These results indicate that for very large systems optimal observation of amide and methyl moieties will require separately labeled samples.

Purification of Arp2/3 Complex Labeled at a Single Subunit. Having established that methyl resonances can, in principle, be observed in very large systems, we sought to extend our analyses to Arp2/3 complex. This 240 kDa assembly is representative of asymmetric, multimeric macromolecular machines. In such systems, spectral complexity must be reduced by selective labeling. Thus, as a first step we have developed a strategy to produce Arp2/3 complex labeled for NMR at an individual, bacterially expressed subunit. This is based on reconstitution of p21 into a hexameric mutant yeast complex (Δyp21) in which the yeast homologue of p21 has been knocked out by homologous recombination. The remaining six subunits remain associated with one another (*29*). p21 interacts with activating WASP peptides in yeast two-hybrid and chemical cross-linking analyses, and the crystal structure of inactive Arp2/3 complex shows that p21 interacts only with Arp3 (*16*). Thus, this subunit is an interesting target for studies of the activation mechanism of Arp2/3 complex. The Δyp21 complex alone is inactive in actin polymerization assays, and unlike the wild-type assembly, it cannot be stimulated by an activating WASP peptide (Figure 4A). Activity of Arp2/3 complex is significantly recovered on the addition of recombinant bacterially expressed human p21, indicating that this subunit is important for WASP-mediated activation of Arp2/3 complex and has incorporated into the mutant complex to restore its activity. We chose to reconstitute Δyp21 with

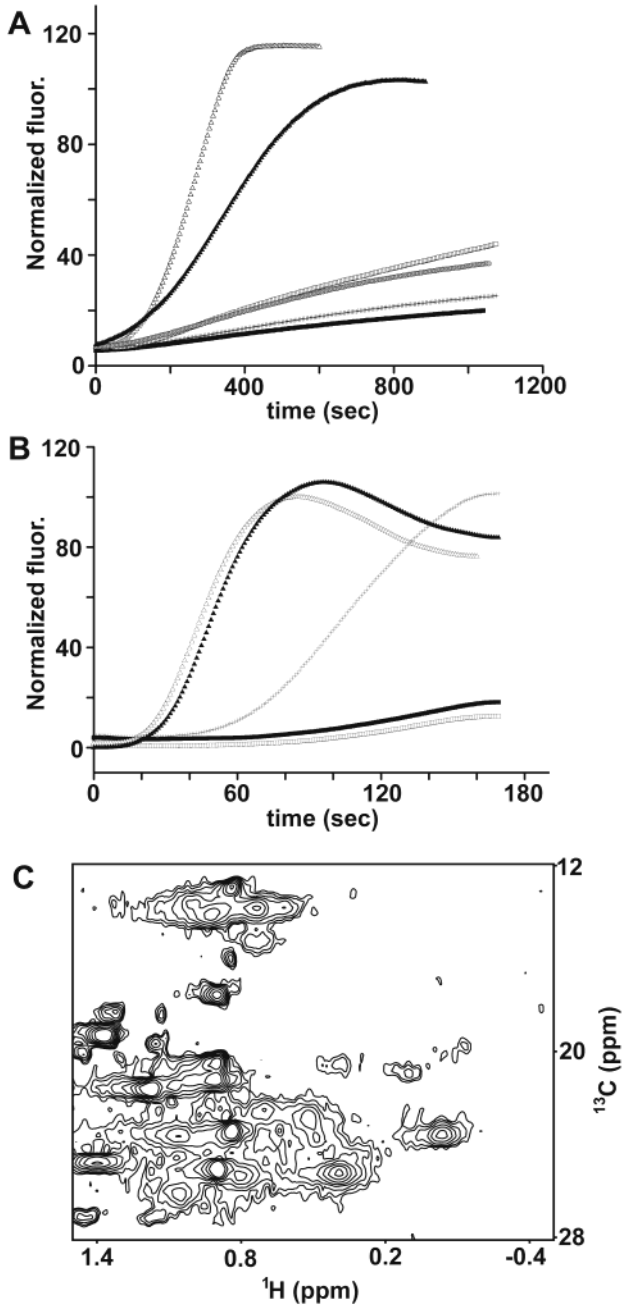

FIGURE 4: Activity and NMR studies of Arp2/3 complex selectively labeled at p21. (A) Recovery of Δyp21 activity upon addition of exogenous p21. Actin (1.5 μM pyrene-actin mix, 10% labeled) was assayed alone (■) or in the presence of 0.125 μM yeast Arp2/3 complex (□), 0.125 μM yeast Arp2/3 + 0.5 μM yeast GST-VCA (△), 0.125 μM Δyp21 (+), 0.125 μM Δyp21 + 0.5 μM yeast GST-VCA (○), or 0.125 μM Δyp21 + 6 μM $(His)_{10}$-p21 + 0.5 μM yeast GST-VCA (▲). (B) Activity of selectively labeled Arp2/3 complex. Actin (2 μM pyrene−actin mix, 5.5% labeled) was assayed alone (■) or in the presence of 15 nM bovine Arp2/3 complex (□), 15 nM bovine Arp2/3 + 1 μM N-WASP VCA (△), 15 nM reconstituted Arp2/3 complex (+), or 15 nM reconstituted Arp2/3 complex + 1 μM N-WASP VCA (▲). (C) ^{1}H/^{13}C-HSQC spectrum of reconstituted Arp2/3 complex ^{13}C/^{15}N/^{2}H/^{1}H(methyl)-labeled at p21. Spectrum was recorded in 96 h at 25 °C at 600 MHz.

human p21 because yeast p21 is unstable and tends to aggregate in solution, and the homology between the two proteins is high enough (46% identical) that human p21 should associate with Δyp21 with reasonable affinity. We have used this ability of p21 to incorporate into Δyp21 to generate an Arp2/3 complex labeled only at p21 (see Experimental Procedures). Briefly, ^{13}C,^{15}N,^{2}H-labeled p21 selectively protonated at the side chain methyl groups of Leu, Val, and Ile (δ_1 only) residues (*9*) was added to the cleared lysate from 1 kg (~300 L culture) of Δyp21 yeast cells. The reconstituted complex (^{13}C,^{15}N,^{2}H,^{1}H(methyl) at p21 and ^{12}C,^{14}N,^{1}H at all other subunits) was then isolated on IgG Sepharose using a Protein A tag conjugated to Arp3, followed by Tev protease cleavage of the tag, cation exchange, and gel filtration chromatographies. This results in a final sample concentration of ~100 μM in a volume of 250 μL. This reconstituted assembly, on stimulation by an activating peptide from N-WASP, has activity comparable to that of wild-type bovine Arp2/3 complex (Figure 4B).

^{1}H/^{13}C-HSQC Spectrum of Arp2/3 Complex. We acquired a ^{1}H/^{13}C-HSQC spectrum (96 h acquisition time) of selectively labeled Arp2/3 complex at 25 °C at 600 MHz (Figure 4C). Of the 51 Leu, Val, and Ile methyl groups in p21, signals from about half of them are visible in Arp2/3 complex under these conditions. Nine of the 11 Ile δ1 methyls in p21 can be observed, consistent with the observation that Ile methyl signals are typically more intense than those from Val and Leu side chains. The methyl resonances not detected are likely broadened by a combination of exchange effects and proximity to other methyl-containing residues. Because the selectively labeled complex is a hybrid assembly of yeast and human proteins, the affinity of human p21 for yeast Arp3 is likely reduced relative to that of yeast p21. This may result in exchange broadening of p21 resonances located near the p21−Arp3 interface. The reason for little effect on the number of Ile signals that are detectable may be that most Ile signals are intense enough to tolerate broadening, while Val and Leu signals are not. The two Ile signals that are not detectable in this spectrum could be lost due to their proximity to multiple methyl-containing residues within p21. Analysis of the crystal structure of Arp2/3 complex (*16*) shows that while most Ile δ1 carbons have zero to two methyl carbons within 5 Å of them, two residues (Ile 41 and Ile 71) are within 5 Å of three and four methyl carbons, respectively. The δ1 methyl protons of these two side chains, then, will be subject to dipolar interactions with more nearby protons than those of the other nine Ile residues, and this may result in much more rapid relaxation of these nuclei. While some p21 signals appear to be severely broadened, the remaining signals in the spectrum will be useful probes in subsequent studies of interactions of Arp2/3 complex with activating WASP peptides.

DISCUSSION

We have explored the utility of methyl groups as NMR probes in large systems. We find that in a perdeuterated background, protonated side chain methyl groups on Leu, Val, and Ile residues produce detectable signals in ^{1}H/^{13}C-HSQC spectra for molecules with correlation times as long as about 330 ns. Analysis of the sensitivity and ^{1}H line widths as a function of τ_C in the ^{1}H/^{13}C-HSQC spectra of GDIΔ59 reveals interesting technical benefits of studying methyl groups in large biomolecules. Because of the high sensitivity and narrow lines seen for Ile signals, the acquisition time required to obtain reasonable S/N in studies of the 0.8 mM GDIΔ59 sample used here was only 4.5 h. This indicates that reasonably short acquisition times can be used, even

for dilute samples, to detect methyl signals from Ile residues. This is already apparent for Arp2/3 complex, where only 96 h were required to detect methyl signals from a ~100 μM sample. These times will be made even shorter as a result of newly available cryogenic probe technology, which will increase sensitivity 2−3-fold, greatly expanding the set of macromolecular systems amenable to NMR analysis. It is also interesting to note that the sensitivity in $^1H/^{13}C$-HSQC spectra is actually higher at 600 MHz than 800 MHz for long τ_C, presumably because of slower proton T_1 relaxation at the higher field. The acquisition of high-quality spectra of methyl resonances in large biological systems, then, does not require the use of the higher-field NMR spectrometers (800 and 900 MHz) needed for optimal performance of TROSY-based sequences used to study amide groups in large molecules. Another advantage of methyl groups over amides is that in p21 and GDIΔ59, many amide sites remain deuterated from bacterial growth because of slow exchange with H_2O solvent during purification. These sites, of course, are not uniformly distributed. While amides can remain deuterated in a structure-dependent fashion after production in D_2O media, methyls are uniformly protonated to high levels. Therefore, they can be evenly observed throughout the structure of a protein, and the absence of signals (or variations in intensity) can be interpreted in a more straightforward manner.

We have also shown that protonation of side chain methyl groups lowers the intensity of amide signals observed using TROSY-based experiments, presumably because of interproton dipolar interactions. Thus, optimal performance of the TROSY sequences cannot be achieved in methyl-labeled samples. While TROSY-type experiments for amides and methyl $^1H/^{13}C$ correlation experiments offer complimentary methods to study the details of conformational changes and ligand interactions in very large systems, effective simultaneous use of both techniques will require separate, appropriately labeled samples.

While methyl resonances can be observed from Ile, Leu, and Val residues in very large systems, signals from Ile δ1 methyl groups are consistently more intense. The weaker intensity of Leu and Val signals is most likely due to dipolar interactions between diastereotopic methyl protons in these side chains. This could potentially be remedied by selective protonation at only one of the two diastereotopic methyl groups. We have recently explored this possibility by labeling with 2-keto-3-methyl-d_3-butyric-4-^{13}C acid now available through certain manufacturers. Incorporation of this precursor into Val and Leu results in the desired pattern, where for each residue in any given protein molecule only one of the two methyl groups is protonated. However, since this precursor is achiral, labeling is nonstereospecific, and the population is protonated to only 50%. It is not yet clear if the 2-fold decrease in labeled population is compensated by improved relaxation properties, although this should be the case as molecular weight increases and relaxation effects become more severe. It may ultimately be possible to synthesize stereoselectively labeled Val, which could lead to stereoselective labeling of this residue in proteins. However, current strategies for labeling Leu, involving biosynthesis from labeled Val (*30*−*32*) or α-ketoisovaleric acid (*9*), would not work here since the biosynthetic pathway leads to scrambling of the methyl label. Thus, stereoselective Leu labeling will likely require stereoselective synthesis of this residue.

NMR analysis of large, asymmetric, multimeric assemblies requires generation of the systems in selectively labeled form. In many cases it will be necessary to purify some of the components, either individually or as a sub-assembly, from a robust eukaryotic expression system such as yeast or insect cells and then introduce labeled component(s) expressed separately in bacteria. We were forced to develop such a purification strategy for Arp2/3 complex labeled for NMR at the p21 subunit since most of its subunits cannot be expressed in bacteria. An inherent drawback to such an approach is that since eukaryotic cells are difficult or prohibitively expensive to grow in D_2O, the majority of the assembly is necessarily generated in unlabeled form (i.e., ^1H). This could lead to significant relaxation in the labeled component brought about by $^1H-^1H$ dipolar interactions with protons on its binding partner. While this does not appear to be a problem in CRIPT/TROSY analysis of amides (*8*), the closer proximity of methyl side chains to protonated groups across an interface suggests that it may be a larger issue in the approach here. Dynamics at subunit interfaces may be an additional source of relaxation in such systems. Although Arp2/3 complex appears to be a constitutive assembly in vivo, the energetics (and therefore dynamics) at its subunit interfaces are unknown. It is possible that its functional behavior, involving changes between multiple tertiary and quaternary conformations, may require significant dynamics within the assembly. While these dynamics could be essential for function of the assembly, they could significantly degrade the quality of its spectra. An additional difficulty in reconstitution is that due to the biochemical behavior of individual subunits, it may also be necessary (as it is for Arp2/3 complex) to reconstitute macromolecular assemblies with components from different organisms. The incorporation of human p21 into yeast Arp2/3 complex, for example, likely enhances intersubunit broadening effects because of reduced affinity between subunits in the chimeric assembly. These combined effects likely contribute to the observed broadening of p21 resonances at the Arp3 interface in $^1H/^{13}C$ HSQC spectra. In cases where structure of an assembly was not known, these could be used to identify interfaces within the assembly.

The overall quality of Arp2/3 complex spectra is lower than expected based on molecular weight alone (e.g. compare Figures 4 and 1 C, G). This may result from nonisotropic tumbling of the assembly, due to its disklike structure. Alternatively, we note that the spectra of free p21 are much poorer than expected for an ideal protein of 21 kDa (not shown). In contrast, spectra of GDIΔ59 are outstanding for a protein of 16 kDa. Such differences in spectral quality between proteins of comparable molecular weight are widely observed. Although not well-understood, they are generally ascribed to differences in chemical exchange dynamics. The properties of free p21 that degrade the quality of its spectra relative to GDIΔ59 may simply persist in reconstituted Arp2/3 complex. Further studies of other very large systems are needed to learn if this is typical of multimeric assemblies, or specific to Arp2/3 complex.

Obtaining chemical shift assignments in very large systems remains a significant challenge. TROSY-based triple-resonance strategies have been successfully used to obtain

backbone assignments in both symmetric and asymmetric systems up to approximately 110 kDa. While these strategies are likely to have only limited utility in appreciably larger systems (*2*), their extension to CRINEPT-based methods may be successful, particularly with the advent of increasingly higher-field magnets. However, even simple $^{1}H/^{15}N-$CRINEPT/TROSY spectra of Arp2/3 complex, or of GDIΔ59 under conditions where τ_C is greater than $\sim$ 150 ns, are not of high quality (not shown), indicating that more complex experiments needed for backbone assignment are unlikely to be successful. Assignment may be more difficult for methyl groups than for amides. Traditional strategies used to obtain methyl chemical shift assignments involve relaying methyl magnetization to previously assigned resonances in the backbone (e.g., HN (*33*, *34*), Ha (*35*), or side chain (*35*−*37*)). Very recently, experiments of the former class were used to assign Ile δ1 resonances in malate synthase G, an 81.4 kDa protein with $\tau_C \approx$ 46 ns (*38*) and to assign Val, Leu, and Ile methyl resonances in the integral membrane protein OmpX solubilized in detergent micelles, with a total molecular weight of ~80 kDa (*33*). However, such strategies will fail when dynamic properties are such that backbone assignments cannot be obtained. Thus, new strategies must be developed to assign methyl resonances in very large systems. In systems such as Arp2/3 complex, where an individual component can be assigned in isolation, many chemical shift assignments of the subunit in the intact assembly may be directly mapped from those of the free protein. In fact, we have obtained tentative assignments for many methyl groups in p21 in Arp2/3 complex in this manner (not shown). We note that while intersubunit contacts will perturb resonances at the interface (making them harder to assign by mapping), we expect that residues involved in ligand binding and thus of significant interest will not be buried at interfaces and thus will be more easily mapped. In specific cases, it may be possible to assign some resonances by mutagenesis, although this approach suffers from the need to prepare numerous samples, and from potential effects on folding and function. The availability of a high-resolution structure of a large assembly may enable use of a more novel strategy involving analysis of pseudo-contact shifts induced by paramagnetic agents incorporated at specific known sites to determine assignments. This strategy had been used before the advent of triple-resonance methods to assign ^{13}C chemical shifts in reductively methylated lysine side chains in concanavalin A (*39*). An analogous strategy was also recently used to resolve symmetry-related overlapped resonances in a symmetric dimer of STAT-4 (*40*). An advantage of a pseudo-contact shift-based approach is that it relies only on changes in chemical shift, a property that can be measured with high sensitivity and accuracy. In summary, while it may be possible to obtain chemical shift assignments of backbone and methyl groups through CRINEPT-based extension of current resonance correlation strategies, alternative approaches are likely to be necessary for applications to very large systems.

Many of the key players and intermediates along biological pathways are very large proteins and oligomeric assemblies. A complete understanding of the mechanisms of action of these molecules will require detailed biochemical and biophysical studies of their structures, dynamics, and ligand interactions. Because methyl-containing residues are enriched in protein cores and at protein−protein binding interfaces, they will be useful probes of ligand interactions and conformational changes in large systems. A variety of techniques can be applied to methyl groups to monitor such events, including chemical shift perturbation, paramagnetic line broadening and pseudocontact shifts, and intermolecular NOEs. NMR spectroscopy of side chain methyl groups, alone or in concert with TROSY-based studies of amides, will prove to be an extremely powerful technique to study important aspects of macromolecular function.

NOTE ADDED IN PROOF

We have recently obtained higher quality data on Arp2/3 complex by recording hmqc spectra, which Kay and coworkers (*47*) have elegantly demonstrated has an inherent TROSY effect because of intramethyl dipolar interactions.

ACKNOWLEDGMENT

We thank Lewis Kay for providing the CRINEPT pulse sequence used in this work, sharing preliminary data, and initiating single methyl group labeling in Val and Leu. Part of the NMR data was acquired at the Environmental Molecular Sciences Laboratory (a national scientific user facility sponsored by the United States DOE Office of Biological and Environmental Research) located at Pacific Northwest National Laboratory, operated by Battelle for the DOE.

SUPPORTING INFORMATION AVAILABLE

Figure depicting correlation of τ_C with protein molecular weight. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−71.
2. Riek, R., Wider, G., Pervushin, K., and Wuthrich, K. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 4918−23.
3. Fernandez, C., Adeishvili, K., and Wuthrich, K. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 2358−63.
4. Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (2000) *J. Am. Chem. Soc. 122*, 7543−8.
5. Evenas, J., Mittermaier, A., Yang, D., and Kay, L. E. (2001) *J. Am. Chem. Soc. 123*, 2858−64.
6. Yang, D., and Kay, L. E. (1999) *J. Am. Chem. Soc. 121*, 2571−5.
7. Tugarinov, V., Muhandiram, D. R., Ayed, A., and Kay, L. E. (2002) *J. Am. Chem. Soc. 124*, 10025−35.
8. Fiaux, J., Bertelsen, E. B., Horwich, A. L., and Wuthrich, K. (2002) *Nature 418*, 207−11.
9. Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. (1999) *J. Biomol. NMR 13*, 369−74.
10. Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T., and Kay, L. E. (1996) *J. Mol. Biol. 263*, 627−36.
11. Kay, L. E., Bull, T. E., Nicholson, L. K., Griesinger, C., Schwalbe, H., and Torchia, D. A. (1992) *J. Magn Reson. 100*, 538−58.
12. Kay, L. E., and Gardner, K. H. (1997) *Curr. Opin. Struct. Biol. 7*, 722−31.
13. Hajduk, P. J., Augeri, D. J., Mack, J., Mendoza, R., Yang, J., Betz, S. F., and Fesik, S. W. (2000) *J. Am. Chem. Soc. 122*, 7898−904.
14. Janin, J., Miller, S., and Chothia, C. (1988) *J. Mol. Biol. 204*, 155−64.
15. Pellecchia, M., Meininger, D., Dong, Q., Chang, E., Jack, R., and Sem, D. S. (2002) *J. Biomol. NMR 22*, 165−73.
16. Robinson, R. C., Turbedsky, K., Kaiser, D. A., Marchand, J. B., Higgs, H. N., Choe, S., and Pollard, T. D. (2001) *Science 294*, 1679−84.

17. Volkmann, N., Amann, K. J., Stoilova-McPhie, S., Egile, C., Winter, D. C., Hazelwood, L., Heuser, J. E., Li, R., Pollard, T. D., and Hanein, D. (2001) *Science 293*, 2456−9.
18. Gosser, Y. Q., Nomanbhoy, T. K., Aghazadeh, B., Manor, D., Combs, C., Cerione, R. A., and Rosen, M. K. (1997) *Nature 387*, 814−9.
19. Winter, D., Podtelejnikov, A. V., Mann, M., and Li, R. (1997) *Curr. Biol. 7*, 519−29.
20. Winter, D., Lechler, T., and Li, R. (1999) *Curr. Biol. 9*, 501−4.
21. Higgs, H. N., Blanchoin, L., and Pollard, T. D. (1999) *Biochemistry 38*, 15212−22.
22. Pollard, T. D., and Cooper, J. A. (1984) *Biochemistry 23*, 6631−41.
23. Spudich, J. A., and Watt, S. (1971) *J. Biol. Chem. 246*, 4866−71.
24. Cooper, J. A., Walker, S. B., and Pollard, T. D. (1983) *J. Muscle Res. Cell Motil. 4*, 253−62.
25. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
26. Sheely, M. L. (1932) *Ind. Eng. Chem. 24*, 1060−4.
27. Green, E., and Parke, J. P. (1939) *J. Soc. Chem. Ind. 58*, 319−20.
28. Wand, A. J., Urbauer, J. L., McEvoy, R. P., and Bieber, R. J. (1996) *Biochemistry 35*, 6116−25.
29. Winter, D. C., Choe, E. Y., and Li, R. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 7288−93.
30. Gardner, K. H., and Kay, L. E. (1997) *J. Am. Chem. Soc. 119*, 7599−600.
31. Smith, B. O., Ito, Y., Raine, A., Teichmann, S., Ben-Tovim, L., Nietlispach, D., Broadhurst, R. W., Terada, T., Kelly, M., Oschkinat, H., Shibata, T., Yokoyama, S., and Laue, E. D. (1996) *J. Biomol. NMR 8*, 360−8.
32. Metzler, W. J., Wittekind, M., Goldfarb, V., Mueller, L., and Farmer, B. T. (1996) *J. Am. Chem. Soc. 118*, 6800−1.
33. Hilty, C., Fernandez, C., Wider, G., and Wuthrich, K. (2002) *J. Biomol. NMR 23*, 289−301.
34. Lohr, F., and Ruterjans, H. (2002) *J. Magn. Reson. 156*, 10−18.
35. Uhrin, D., Uhrinova, S. C. L., Nairn, J., Price, N. C., and Barlow, P. N. (2000) *J. Magn. Reson. 142*, 288−93.
36. Hu, W., and Zuiderweg, E. R. P. (1996) *J. Magn. Reson. B 113*, 70−5.
37. Zwahlen, C., Vincent, S. J. F., Gardner, K. H., and Kay, L. E. (1998) *J. Am. Chem. Soc. 120*, 4825−31.
38. Tugarinov, V., and Kay, L. E. (2003) *J. Am. Chem. Soc. 125*, 5701−6.
39. Sherry, A. D., and Teherani, J. (1983) *J. Biol. Chem. 258*, 8663−9.
40. Gaponenko, V., Altieri, A. S., Li, J., and Byrd, R. A. (2002) *J. Biomol. NMR 24*, 143−8.
41. Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York.
42. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego.
43. Creighton, T. E. (1984) *Proteins: Structures and Molecular Properties*, W. H. Freeman & Co., New York.
44. Yguerabide, J., Epstein, H. F., and Stryer, L. (1970) *J. Mol. Biol. 51*, 573−90.
45. Krause, S., and O'Konski, C. T. (1963) *Biopolymers 1*, 503−15.
46. *CRC Handbook of Chemistry and Physics*, 83rd ed. (2002) CRC Press, Boca Raton.
47. Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E., and Kay, L. E. (2003) Cross-Correlated Relaxation Enhanced ^{1}H-^{13}C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes, submitted.

BI034536J