Global Folds of Highly Deuterated, Methyl-Protonated Proteins by Multidimensional NMR†

Kevin H. Gardner,* Michael K. Rosen,‡ and Lewis E. Kay

Protein Engineering Network Centres of Excellence and Departments of Medical Genetics and Microbiology, Biochemistry, and Chemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Received October 2, 1996; Revised Manuscript Received November 27, 1996⊗

ABSTRACT: The development of ^{15}N, ^{13}C, ^{2}H multidimensional NMR spectroscopy has facilitated the assignment of backbone and side chain resonances of proteins and protein complexes with molecular masses of over 30 kDa. The success of these methods has been achieved through the production of highly deuterated proteins; replacing carbon-bound protons with deuterons significantly improves the sensitivity of many of the experiments used in chemical shift assignment. Unfortunately, uniform deuteration also radically depletes the number of interproton distance restraints available for structure determination, degrading the quality of the resulting structures. Here we describe an approach for improving the precision and accuracy of global folds determined from highly deuterated proteins through the use of deuterated, selectively methyl-protonated samples. This labeling profile maintains the efficiency of triple-resonance NMR experiments while retaining a sufficient number of protons at locations where they can be used to establish NOE-based contacts between different elements of secondary structure. We evaluate how this deuteration scheme affects the sensitivity and resolution of experiments used to assign ^{15}N, ^{13}C, and ^{1}H chemical shifts and interproton NOEs. This approach is tested experimentally on a 14 kDa SH2/phosphopeptide complex, and a global protein fold is obtained from a set of methyl−methyl, methyl−NH, and NH−NH distance restraints. We demonstrate that the inclusion of methyl−NH and methyl−methyl distance restraints greatly improves the precision and accuracy of structures relative to those generated with only NH−NH distance restraints. Finally, we examine the general applicability of this approach by determining the structures of several proteins with molecular masses of up to 40 kDa from simulated distance and dihedral angle restraint tables.

The development of multinuclear, multidimensional NMR[1] spectroscopy has had a significant impact on the utility of NMR in structural biology. Most applications to date have focused on macromolecules with molecular masses of under 25 kDa (Clore & Gronenborn, 1994). However, it is clear that deuteration of aliphatic positions in ^{15}N, ^{13}C-labeled samples can further increase the molecular mass of systems amenable to investigation by NMR. The use of deuteration for spectral improvement and simplification dates back to experiments performed in the late 1960s by Markley and Jardetzky as well as by Crespi et al. (Crespi et al., 1968; Markley et al., 1968; Crespi & Katz, 1969). Subsequently, several groups demonstrated the utility of fractional or complete deuteration in 2D homonuclear spectroscopy (LeMaster & Richards, 1988; Torchia et al., 1988; Tsang et al., 1990), and Jardetzky, Arrowsmith, and co-workers used a scheme combining deuteration and heteronuclear-edited NMR methods to determine the solution structure of a 37 kDa *trp* repressor protein/DNA complex (Zhang et al., 1994).

Perhaps the most important applications of deuteration are in combination with triple-resonance (^{15}N, ^{13}C, ^{1}H) spectroscopy. Experiments recorded on ^{15}N, ^{13}C, ^{2}H-labeled samples demonstrate significantly improved spectral resolution and sensitivity relative to similar experiments performed on fully protonated molecules (Grzesiek et al., 1993; Yamazaki et al., 1994a,b; Farmer & Venters, 1995; Nietlispach et al., 1996). These benefits arise from the lower gyromagnetic ratio of deuterons relative to that of protons ($\gamma_H/\gamma_D \sim 6.5$), substantially lengthening the relaxation times of carbon and proton spins in proximity to the substituted deuterons. With the development of experiments to take advantage of the favorable relaxation properties of ^{2}H-bound carbon nuclei, deuterated samples are becoming more commonly used in NMR studies of proteins with molecular

† This work was supported through grants from the Natural Sciences and Engineering Research Council of Canada, the Medical Research Council of Canada, and the National Cancer Institute of Canada with funds from the Canadian Cancer Society. K.H.G. and M.K.R. were supported by postdoctoral fellowships from the Helen Hay Whitney Foundation and the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation, respectively.

* Corresponding author.

‡ Current address: Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021.

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1997.

[1] Abbreviations: 2D, two-dimensional; 3D, three-dimensional; 4D, four-dimensional; CSI, chemical shift index; CT, constant time; DG/SA, distance geometry/simulated annealing; (HB)CBCA(CO)NH, α/β carbon (via carbonyl carbon) to nitrogen to amide proton correlation; (H)C(CO)NH-TOCSY, side chain carbon (via carbonyl carbon) to nitrogen to amide proton correlation; HNCO, amide proton to nitrogen to carbonyl carbon correlation; HN(CO)CA, amide proton to nitrogen (via carbonyl carbon) to α-carbon correlation; HN(COCA)CB, amide proton to nitrogen (via carbonyl carbon and α-carbon) to β-carbon correlation; HSQC, heteronuclear single-quantum coherence; GM-CSF, granulocyte-macrophage colony-stimulating factor; ISPA, isolated spin pair approximation; NOE, nuclear Overhauser effect; NMR, nuclear magnetic resonance; PLCC SH2, C-terminal SH2 domain of bovine phospholipase Cγ1; PLCC/pY1021, complex of the C-terminal SH2 domain of bovine phospholipase Cγ1 and a 12-residue phosphotyrosine peptide comprising the Tyr 1021 site of the platelet-derived growth factor receptor; rf, radio frequency; rmsd, root mean-squared difference of atomic positions; SNase, staphylococcal nuclease.

S0006-2960(96)02480-4 CCC: $14.00

masses of 20 kDa or greater. Two recent examples of this are the structure determinations of the Shc phosphotyrosine binding domain (190 residues) and Bcl-X_L (209 residues), where Fesik and co-workers used 70% random-fractionally deuterated samples (Zhou et al., 1995; Muchmore et al., 1996).

As the molecular mass of the macromolecule increases, so too does the efficiency of dipolar relaxation resulting from the remaining protons. In this context, in a recent study of a 64 kDa *trp* repressor/DNA complex, it was necessary to deuterate the aliphatic positions to a high level (>90%) for the assignment of backbone NH, ^{15}N, ^{13}Cα, and ^{13}Cβ chemical shifts (Shan et al., 1996). It is likely that a similarly high level of ^{2}H incorporation will be needed for the sequential assignment of many single polypeptide chain proteins with molecular masses on the order of 30−40 kDa. This near-complete substitution of deuterons for protons presents a serious problem for NMR-based structure determination with its heavy reliance on interproton distance restraints established by nuclear Overhauser effect (NOE) experiments. Note that the exchangeable NH sites are fully protonated in these highly deuterated samples and distance restraints involving these protons can be obtained using recently developed 4D ^{15}N, ^{15}N-edited NOESY experiments (Grzesiek et al., 1995; Venters et al., 1995). However, simulations and experiments in our laboratory suggest that distance information from backbone NH protons alone will generally be insufficient for obtaining accurate global folds of proteins.

With these problems in mind, we have sought to develop an approach that maximizes the extent of deuteration in a given sample while introducing protons at high occupancy into a small number of specific sites. Samples labeled in this manner should retain many of the benefits that deuteration provides for the NMR experiments used in chemical shift assignment without the concomitant loss of the interproton distance restraints necessary for defining protein global folds. To this end, we have developed a biosynthetic method where most aliphatic positions are deuterated to high levels while the methyl groups of several hydrophobic amino acids remain highly protonated (Rosen et al., 1996). This is achieved by overexpressing protein in bacteria grown in D_2O with protonated [^{13}C]pyruvate as the sole carbon source. Protons attached to the pyruvate methyl carbon are retained in the methyl groups of alanine, valine, leucine, and isoleucine (γ2 methyl position only). Proteins generated in this manner show 40−80% protonation at the pyruvate-derived methyl groups and lower levels of protonation at several methylene locations (Hβ of Ser, Trp, and Asn/Asp; Hγ of Arg, Gln/Glu, and Pro) in an otherwise highly deuterated system.

Methyl groups are excellent targets for protonation from an NMR standpoint as ^{13}C−^{1}H correlation spectra are often well-resolved and fast rotation about the methyl symmetry axis leads to relatively narrow proton and carbon spectral line widths (Kay et al., 1992a). Furthermore, Ala, Val, Ile, and Leu are the most common amino acids in protein interiors (Janin et al., 1988). Therefore, protonated methyl groups from these residues are often within 5 Å (NOE distance) of other residues on different secondary structural elements. These groups are also highly represented at molecular interfaces, suggesting that this method may also be useful for studies of protein complexes.

Protonated methyl groups offer an important source of structural information in deuterated proteins; here we evaluate the significance of this contribution in the process of protein structure determination. We initially examine how the presence of protonated methyl groups impacts the assignment of chemical shifts, in terms of both experimental sensitivity and deuterium isotope effects on the observed chemical shifts. Experimental NOE data from a 14 kDa complex of an ^{15}N, ^{13}C, ^{2}H-labeled SH2 domain with protonated methyl groups of Ala, Val, Leu, and Ile (γ2 only) [referred to as an ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl)-labeled SH2 domain] bound to an unlabeled 12-residue phosphopeptide are used to compile NH−NH, methyl−NH, and methyl−methyl distance restraints. Structures generated from this data together with dihedral angle restraints derived from carbon chemical shifts are presented, demonstrating that the addition of distance restraints involving methyl groups significantly improves both the accuracy and precision of the resulting structures. The generality of this approach for application to larger proteins is investigated by determining structures from simulated sets of dihedral angle and distance restraints derived from several protein systems up to 40 kDa in molecular mass. These analyses demonstrate that the structural information available from ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl)-labeled samples will be sufficient in many cases to determine the global folds of larger proteins.

MATERIALS AND METHODS

Sample Preparation. An ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl)-labeled sample of the bovine phospholipase Cγ1 (PLCγ1) C-terminal SH2 domain (PLCC SH2) was prepared as described previously (Rosen et al., 1996). This protein contains PLCγ1 amino acid residues 663−759 and additional residues introduced during cloning to bring the total length to 105 amino acids. The protein was generated by overexpression in *Escherichia coli* strain BL21(DE3) in M9 media with 99.9% D_2O, $^{15}NH_4Cl$, and [^{13}C]pyruvate and subsequently purified as before. The sample consisted of 0.5 mL of 2 mM PLCC SH2 contained in argon-purged 100 mM sodium phosphate (pH 6.0) buffer with 100 μM DTT-d_{10}, 100 μM EDTA-d_{12}, and 90% H_2O/10% D_2O. The PLCC SH2/pY1021 complex was generated by titrating a 12-residue phosphotyrosine-containing peptide derived from the Tyr 1021 site of the platelet-derived growth factor receptor into the PLCC SH2 sample. The progress of the titration was monitored by observing chemical shift changes in ^{15}N−^{1}H HSQC spectra of the PLCC SH2. A fully protonated, ^{15}N, ^{13}C-labeled PLCC SH2 sample was similarly prepared to measure the effects of ^{2}H substitution on ^{15}N, ^{13}C, and ^{1}H chemical shifts.

Nuclear Magnetic Resonance Experiments. NMR spectra were acquired using a Varian Unity+ 600 MHz spectrometer equipped with a 5 mm triple-resonance probe with a pulsed-field, actively shielded *z*-gradient and a gradient amplifier unit. Unless otherwise noted, experiments were conducted at 30 °C using samples in 90% H_2O/10% D_2O. Spectra were processed and analyzed using the software packages NMRPipe (Delaglio et al., 1995) and NMRView (Johnson & Blevins, 1994), respectively.

Backbone ^{15}N, NH, ^{13}Cα, ^{13}Cβ, and ^{13}CO assignments of the deuterated, methyl-protonated PLCC/pY1021 complex were obtained using constant-time (CT) HN(CO)CA (Yamazaki et al., 1994a), HN(COCA)CB (Shan et al., 1996), and

HNCO (Kay et al., 1994) experiments. The former two are methods that are designed for use on highly deuterated proteins. Methyl ^{13}C resonances were assigned with data from a modified (H)C(CO)NH-TOCSY experiment (Gardner et al., 1996) that correlates these chemical shifts with those of the backbone ^{15}N and NH of the following residue. To establish the utility of this method on high-molecular mass proteins, data were recorded at 3°C; at this temperature, the PLCC/pY1021 complex has a correlation time of 18.8 ± 0.6 ns as determined by ^{15}N relaxation experiments (Farrow et al., 1994). Valine and leucine methyl groups were stereospecifically assigned by comparison to the chemical shifts of the fully protonated PLCC/pY1021 complex (Muhandiram et al., 1995); de novo assignments are readily obtained using a previously established method (Neri et al., 1988). Backbone ^{15}N, NH, and ^{13}C (Cα, Cβ, and CO) chemical shifts of an ^{15}N, ^{13}C fully protonated PLCC/pY1021 sample were obtained from HNCO and (HB)CBCA(CO)-NH experiments (Grzesiek & Bax, 1992; Muhandiram & Kay, 1994), making use of previously determined chemical shifts (Pascal et al., 1994; S. M. Pascal, A. U. Singer, L. E. Kay, and J. D. Forman-Kay, manuscript in preparation).

Distance restraints were determined with data obtained from three 4D gradient-enhanced, isotope-edited NOESY experiments: ^{13}C, ^{13}C-edited (Vuister et al., 1993), ^{13}C, ^{15}N-edited (Muhandiram et al., 1993), and ^{15}N, ^{15}N-edited (Grzesiek et al., 1995; Venters et al., 1995). The two ^{15}N-edited experiments incorporate sensitivity-enhanced, gradient-based coherence transfer pathway selection (Kay et al., 1992b; Schleucher et al., 1993), and care was taken to minimize the perturbation of the H_2O resonance through the use of water-selective flipback pulses (Grzesiek & Bax, 1993; Kay et al., 1994; Stonehouse et al., 1994). Water suppression in the ^{13}C, ^{13}C-edited NOESY experiment was achieved on the basis of two dephasing pulses with orthogonal phases (durations of 6.0 and 3.7 ms; 7.8 kHz rf field strength) sequentially applied before the recycle delay and an H_2O-selective 90° pulse followed immediately by a gradient applied at the end of the NOE mixing period. Each experiment required approximately 5 days of spectrometer time using 4 scans/FID, a 1 s recycle time, a 175 ms NOE mixing period, and a T of 30 °C. The sweep widths and maximum evolution times for each dimension of these experiments were as follows [t_1 (^{1}H), t_2 (X), t_3 (X), and t_4 (^{1}H)]: ^{13}C, ^{13}C-edited NOESY, 3600 Hz (16.7 ms), 3600 Hz (4.4 ms), 3600 Hz (3.3 ms), and 7199 Hz (48.9 ms); ^{13}C, ^{15}N-edited NOESY, 3600 Hz (13.9 ms), 3600 Hz (^{13}C, 4.4 ms), 1335 Hz (^{15}N, 12.0 ms), and 8500 Hz (52.7 ms); and ^{15}N, ^{15}N-edited NOESY, 2200 Hz (12.7 ms), 1335 Hz (19.5 ms), 1335 Hz (12.0 ms), and 8500 Hz (52.7 ms).

PLCC SH2 Structure Determination and Analysis. Recently, Pascal et al. have obtained a refined solution structure of a complex between the C-terminal SH2 domain of phospholipase Cγ1 and a 12-residue tyrosine-phosphorylated peptide derived from the platelet-derived growth factor receptor (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation). The high precision of these structures (0.66 ± 0.11 Å rmsd for backbone atoms between residues 10 and 100) was achieved through the use of 1874 distance and 197 dihedral angle restraints. We have used these structures to examine the relative quality of structures produced from the reduced subset of conformational constraints available from deuterated, methyl-protonated samples.

Distance restraints were obtained by analyzing the 4D NOESY data sets described above in a conservative manner, requiring the unambiguous assignment and identification of both symmetry-related cross-peaks in the ^{15}N, ^{15}N- and ^{13}C, ^{13}C-edited NOESY experiments. Restraints were only tabulated for peaks involving backbone NH and/or methyl protons, although cross-peaks were observed between several other types of protons. These include side chain NH groups and several pyruvate-derived methylene positions, including Hβ for Ser and Asn/Asp and Hγ for Arg, Pro, and Gln/Glu (Rosen et al., 1996). Although not derived directly from pyruvate and protonated at low percentage, several weak NOEs involving Ile Cδ1 methyl groups were observed and used in the structure determination.

Distances were calculated with the isolated spin pair approximation (ISPA). NOE cross-peak intensities were calibrated from the 175 ms mixing time NOESYs using cross-peaks between pairs of atoms separated by known short-range distances that are easily identified in a protein of unknown structure. These atoms and distances are as follows: ^{15}N, ^{15}N-edited NOESY, sequential NH protons in an α-helix (2.5 Å); ^{13}C, ^{15}N-edited NOESY, intraresidue Cβ−NH for Ala in an α-helix (2.6 Å); and ^{13}C, ^{13}C-edited NOESY, intraresidue methyl−methyl distance for Leu and Val (2.5 Å). All restraints based on NOEs to methyl groups were treated as distances to the methyl carbon(s). The pyruvate-based labeling technique leads to residue type-specific differences in methyl protonation efficiency, ranging from 39% (alanine) to 80% (leucine) (Rosen et al., 1996). To account for these differences and their effect on the intensity of NOEs involving methyl groups, the intensities of methyl NOEs were divided by the average number of protons present in each particular methyl type via (Yip, 1990)

$$a'_{AB}(\tau_m) = \frac{a_{AB}(\tau_m)}{n_A n_B} \quad (1)$$

where $a'_{AB}(\tau_m)$ and $a_{AB}(\tau_m)$ are the corrected and observed intensities of the NOESY cross-peak between sites A and B, respectively, and n_A and n_B are the average numbers of protons at A and B, respectively. For example, consider an NOE observed between the methyl group of an alanine residue (39% protonation, $n_A = 1.17$) and a methyl group of a leucine residue (80% protonation, $n_B = 2.40$). The intensity of this NOE is divided by the product of the average number of protons present on each methyl group ($1.17 \times 2.40 = 2.81$) prior to determining the distance via the ISPA.

Distance restraints from the ^{15}N, ^{15}N- and the ^{13}C, ^{15}N-edited NOESY experiments employed a lower bound of 1.8 Å and an upper bound 33% larger than the ISPA-calculated distance. A comparison of the distances obtained from the 4D data sets with distances calculated on the basis of the mean structure obtained from the group of highly refined PLCC SH2 reference structures (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation) indicated that this method overestimated the actual distances by an average of 18% (NH−NH NOEs) and 11% (methyl−NH NOEs). This comparison also suggested that it was not possible to accurately quantitate distances from the ^{13}C, ^{13}C-edited NOE data, likely due to spin diffusion; as such, all upper bounds for these restraints were set to the longest experimentally observed distance (a 7 Å carbon−carbon distance). Although the PLCC SH2 was complexed to a fully protonated phosphopeptide and protein/peptide NOEs were

Table 1: Chemical Shift Differences between the ^{15}N, ^{13}C, ^{2}H (^{1}H-Methyl) Domain and the ^{15}N, ^{13}C, ^{1}H PLCC SH2[a]

amino acid	^{15}N	$^{13}C\alpha$	$^{13}C\beta$[b]
Gly	-0.23 ± 0.10	-0.50 ± 0.02[b]	na
Ile	-0.16 ± 0.12	-0.45 ± 0.04	-0.91 ± 0.07
Thr	-0.15	-0.40	-0.57
Val	-0.23 ± 0.07	-0.49 ± 0.05	-0.77 ± 0.02
Asn	-0.28	-0.34 ± 0.01	-0.61 ± 0.01
Asp	-0.27 ± 0.04	-0.43 ± 0.02	-0.70 ± 0.02
Arg	-0.17 ± 0.14	-0.46 ± 0.06	-0.96 ± 0.05
Cys	-0.49	-0.46	-0.65
Gln	-0.19 ± 0.05	-0.43 ± 0.04	-0.82 ± 0.05
Glu	-0.22 ± 0.07	-0.44 ± 0.09	-0.83 ± 0.06
His	-0.18 ± 0.05	-0.44 ± 0.08	-0.72 ± 0.05
Leu	-0.28 ± 0.06	-0.45 ± 0.05	-1.08 ± 0.04
Lys	-0.22 ± 0.06	-0.40 ± 0.04	-1.04 ± 0.06
Met	-0.29 ± 0.05	-0.46 ± 0.10	-0.94 ± 0.03
Phe	-0.24 ± 0.08	-0.37 ± 0.03	-0.90 ± 0.01
Pro	nd	-0.45 ± 0.01	-0.85 ± 0.05
Ser	-0.20 ± 0.10	-0.43 ± 0.09	-0.72 ± 0.06
Trp	-0.15	-0.34	-0.69
Tyr	-0.27 ± 0.11	-0.43 ± 0.05	-0.86 ± 0.04
Ala	-0.22 ± 0.05	-0.42 ± 0.04	-0.93 ± 0.05

[a] No significant changes were observed for NH (0.00 ± 0.01ppm) or ^{13}CO (-0.01 ± 0.05 ppm) chemical shifts. Negative numbers represent upfield shifts of the indicated nucleus upon deuteration; the values listed are averages (±one standard deviation) except for Thr, Cys, and Trp, where only one measurement was made for each shift. Chemical shifts were obtained from spectra with 0.09 ppm/pt [^{13}C for the ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl) sample], 0.23 ppm/pt (^{13}C for the ^{15}N, ^{13}C, ^{1}H sample), 0.07 ppm/pt (^{15}N), and 0.01 ppm/pt (^{1}H) resolution. nd, not determined; na, not applicable. [b] Shifts reported for methylene groups are the difference between the fully protonated CH_2 and fully deuterated CD_2 forms.

utilized in the structure determination of the fully protonated complex (Pascal et al., 1994; S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation), no attempt was made to incorporate intermolecular restraints in this work.

Additional structural information was derived from backbone ϕ and ψ dihedral angle restraints. These were obtained from the secondary structure assignments as determined from the $^{13}C\alpha$, $^{13}C\beta$, and ^{13}CO chemical shifts using the chemical shift index (CSI) method (Wishart & Sykes, 1994). Because the $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts are moved upfield by the replacement of locally bonded protons with deuterons, we have modified the CSI random coil values for each residue by the average difference in shifts between protonated and deuterated, methyl-protonated samples of the PLCC SH2 (Table 1). α-Helical residues were restrained to a (ϕ,ψ) of ($-70 \pm 50°$, $-50 \pm 50°$), while residues in β-sheet conformations were set to ($-140 \pm 60°$, $130 \pm 90°$).

Structures of the PLCC SH2 were generated using an X-PLOR standard hybrid distance geometry/simulated annealing (DG/SA) protocol (Brünger, 1992). Although the pY1021 peptide was present in the NMR sample, restraints involving peptide protons were not included and the peptide was omitted from the structure calculations. To evaluate the quality of the structures that were generated, a number of statistical analyses were performed using the program MOLMOL (Koradi et al., 1996) on structures that were consistent with the input empirical and experimental constraints (less than two violations of NOE restraints over 0.5 Å and less than one violation of dihedral restraints over 5°). These analyses include precision [root mean-squared difference (rmsd) of the positions of backbone heavy atoms in each member of an experimentally generated group and the corresponding atoms in the mean structure of the group] and accuracy [rmsd of backbone heavy atoms in each member of a group and the corresponding atoms in the mean structure of the PLCC SH2 reference group (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation)].

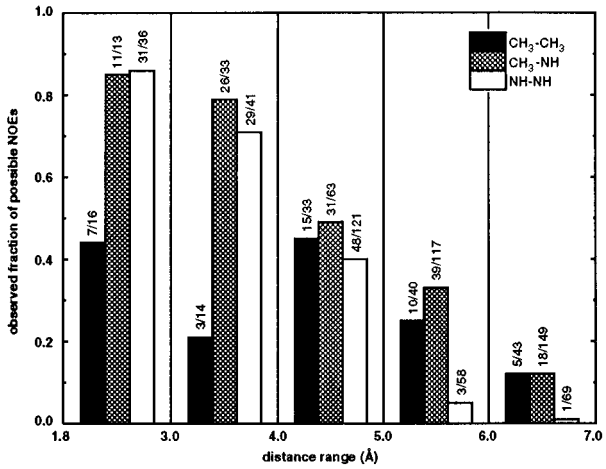

FIGURE 1: Observed fraction of possible NOEs in the PLCC SH2. Possible NOEs were determined from all sets of appropriate protons [or carbons for methyl groups of Ala, Val, Leu, and Ile (γ2 only)] within a given distance range using the mean structure from the set of high-precision reference structures (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation). The numbers of experimentally observed NOEs/possible NOEs are listed above each bar.

Structure Determinations on Larger Proteins Using Simulated Restraints. In order to examine the quality of structures potentially available from larger ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl)-labeled proteins, we generated structures of several different proteins using simulated distance and dihedral angle restraint tables obtained in the following way. Crystallographically determined structures of the following proteins were obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977): granulocyte-macrophage colony-stimulating factor (GM-CSF, 1gmf) (Diederics et al., 1991), staphylococcal nuclease (SNase, 1stn) (Hynes & Fox, 1991), *endo*-1,4-betaxylanase (1xnb) (Campbell et al., 1993), human carbonic anhydrase II (1mua) (B. H. Shilton, H. A. Shuman, and S. L. Mowbray, personal communication), farnesyl diphosphate synthase (1fps) (Tarshis et al., 1994), and maltose binding protein (1mpb) (Tweedy et al., 1993). Structures with a variety of topologies were selected with the MOOSE World Wide Web facility, employing proteins in the 10−45 kDa molecular mass range. Nonprotein components of these structures, such as water and ligand molecules, were removed from all analyses.

Distance restraints were generated by tabulating distances under 7 Å between all combinations of NH and methyl positions that would be protonated in a ^{2}H [Ala, Val, Leu, and Ile (γ2 only), ^{1}H-methyl]-labeled sample. To account for the reduced number of restraints actually observed in experimentally determined sets, restraints from this full list were discarded in a type- and distance-dependent fashion on the basis of their frequency of observation in experimental data sets recorded using the PLCC SH2 complex (Figure 1). Lower and upper bounds were set as in the PLCC SH2 structure determination; all restraints used a 1.8 Å lower bound, while upper bounds were set to 1.18 or 1.11 times the crystallographically determined distance (NH−NH and methyl−NH restraints, respectively) or 7 Å (methyl−methyl

restraints). Note that the values of 1.18 (NH−NH) and 1.11 (methyl−NH) were chosen for consistency with the experimental results from the SH2/pY1021 complex. As described above, it was noted that experimental NH−NH and methyl−NH distances were overestimated by 18 and 11%, respectively, using an upper distance limit set to a factor of 1.33 larger than the distance estimated on the basis of the ISPA. Dihedral angle restraints were determined using the secondary structure assignments in each PDB coordinate file and set to the same ranges as with the PLCC SH2 [α-helix, $(\phi,\psi) = (-70 \pm 50°, -50 \pm 50°)$; and β-sheet, $(\phi,\psi) = (-140 \pm 60°, 130 \pm 90°)$].

These restraints were input into X-PLOR protocols for structure determination. The SNase structure was determined utilizing the same hybrid DG/SA approach employed for the PLCC SH2. The other five structures were determined starting from an extended template using a torsion angle-constrained molecular dynamics approach in X-PLOR (Rice & Brünger, 1994; E. Stein, L. M. Rice, and A. T. Brünger, personal communication). This method offers key advantages over standard molecular dynamics protocols for generating structures of larger proteins as a result of the simpler conformational space that is searched. The searches of torsion angle space are conducted with fixed bond lengths and angles, achieving an approximate 10-fold reduction in the numbers of parameters required to describe the structures. Additionally, the removal of high-frequency bond and angle vibrations allows the use of higher temperatures and longer time steps during refinement, facilitating more computationally efficient searches. The net result is an improved convergence rate and correspondingly shorter times for calculating acceptable structures.

For each protein, five sets of randomly generated distance restraint lists were constructed and combined with the dihedral restraints established as described above; between six and ten structures were obtained from each set of restraints. Subsequently, structures from all five runs were grouped together, and statistics were generated on those that satisfied the angle and the simulated experimental distance restraints. The number of restraint violations that were tolerated in structures used for analysis increased with molecular mass. For example, in the largest system considered (maltose binding protein), the structures had on average 5 NOE violations over 0.5 Å and 15 dihedral angle violations over 5°. For each protein, between 8 and 20 structures were available for analysis.

RESULTS AND DISCUSSION

Chemical Shift Assignment. The initial step in an NMR-based structure determination is the chemical shift assignment of the nuclei required in later analyses. In an ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl)-labeled protein, these include the shifts of backbone (^{15}N, NH, $^{13}C\alpha$, and ^{13}CO) and side chain ($^{13}C\beta$, ^{13}C(methyl), and ^{1}H(methyl)] positions. Triple-resonance experiments optimized for use on large, highly deuterated proteins have been developed for assignment of ^{15}N, NH, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts in proteins and protein complexes (Yamazaki et al., 1994a,b; Shan et al., 1996). These methods rely on magnetization transfer steps through carbon nuclei, requiring high levels of deuteration at these sites for maximal sensitivity. The efficiency of these experiments will remain high when applied to deuterated, methyl-protonated samples due to the significant levels of deuteration at the Cα (102 ± 6%) and Cβ (>75% for all amino acids except Ala, Ser, and Trp) positions (Rosen et al., 1996).

As a prerequisite for the assignment of methyl NOEs, the ^{13}C and ^{1}H chemical shifts of the protonated methyl groups (Ala, Val, Leu, and Ile) must first be obtained. To this end, we have recently developed a 3D (H)C(CO)NH-TOCSY experiment which correlates the ^{13}C(methyl) chemical shift of a residue with the ^{15}N and NH shifts of the following residue (Gardner et al., 1996). The high sensitivity of this experiment is due to the virtually complete level of deuteration in the non-methyl carbon positions of Ala, Val, Leu, and Ile; over 90% of these residues are fully deuterated between the methyl and carbonyl carbon atoms (Rosen et al., 1996). In this way, the transfer of magnetization can proceed along the side chain with minimal relaxation losses. Additionally, the longitudinal relaxation of methyl protons is efficient. An average methyl proton relaxation time T_1 of approximately 0.35 s has been measured for CH_3 groups of the ^{15}N, ^{13}C, ^{2}H-labeled (methyl-protonated) PLCC SH2/pY1021 sample at 3 °C (Rosen et al., 1996). Note that methyl T_1s, unlike NH longitudinal relaxation times, are rather insensitive to field strength and overall molecular correlation time. The efficient relaxation allows the use of short recycle delays between scans in experiments where magnetization originates on methyl protons.

In order to evaluate the (H)C(CO)NH-TOCSY experiment, we have recorded spectra on the ^{15}N, ^{13}C, ^{2}H-labeled (methyl-protonated) PLCC SH2/pY1021 sample at 3 °C. The correlation time measured at this temperature ($\tau_c = 18.8 \pm 0.6$ ns) is similar to that of a 35 kDa globular protein at 30 °C. Virtually all (33/36) of the expected $^{13}CH_3(i)-^{15}N$, NH-$(i+1)$ correlations were observed at high signal to noise, establishing the utility of this method for application to larger proteins. Some correlations were also observed for other protonated sites (Asn/Asp Hβ, Gln/Glu Hγ, and Ser Hβ), but these are at lower intensities primarily due to lower protonation levels and the optimization of several experimental parameters for Ala, Val, Leu, and Ile (Gardner et al., 1996). Modification of the experiment to enable the recording of methyl proton chemical shifts in place of carbon shifts or so both methyl carbon and proton chemical shifts are obtained using a 4D pulse scheme is straightforward.

Effect of Deuteration on Chemical Shifts. The replacement of deuterium for protons leads to an upfield change in the chemical shifts of nearby ^{15}N and ^{13}C nuclei, chiefly via a through-bond mechanism (Hansen, 1988). Studies on small organic compounds have shown that these effects are cumulative; directly bonded deuterons introduce the largest shifts, but significant shifts can be measured due to deuterium incorporation at positions up to four bonds distant from the measured site. In the present case, the magnitudes of the deuterium isotope effects were quantitated by comparing the chemical shifts derived from fully protonated and deuterated, methyl-protonated samples of the PLCC SH2. The amino acid-specific differences for the ^{15}N, $^{13}C\alpha$, and $^{13}C\beta$ shifts are reported in Table 1; deuteration had insignificant effects on the NH and ^{13}CO chemical shifts. Upfield shifts of a fairly uniform magnitude were observed at both the ^{15}N (-0.22 ± 0.09 ppm) and $^{13}C\alpha$ (-0.44 ± 0.08 ppm) positions. From the observed Cα shifts and the extent of Cβ deuteration established in a previous study (Rosen et al., 1996), we estimate the one-bond and two-bond deuterium isotope effects on carbon chemical shifts [$^{1}\Delta$ $^{13}C(^{2}H)$ and $^{2}\Delta$ $^{13}C(^{2}H)$]

to be −0.25 and −0.10 ppm, respectively. These values are quite similar to those observed at aliphatic sites of small organic compounds (Hansen, 1988). A much wider range of shifts is observed at the Cβ position (−0.57 to −1.08 ppm) than at Cα, due to differences in side chain structure and the presence of specifically protonated sites in certain amino acids. Thus, for each amino acid type, there may be differing numbers of protons and deuterons that influence the chemical shift of the ^{13}Cβ carbon. As an illustration, consider the difference in the deuterium-induced shift at the Cβ position of Glu/Gln (−0.83 and −0.82 ppm, respectively) relative to Lys (−1.04 ppm). The Cγ carbons of Glu/Gln are partially protonated as they are derived from a position in α-ketoglutarate that originates from the pyruvate methyl group (Rosen et al., 1996). In contrast, the extent of deuteration at the Cγ position in Lys is much higher. Note that the shifts reported for the methylene Cβ positions (and glycine Cα) are for the fully deuterated CD_2 isotopomers and thus are not affected by differences in protonation at these particular sites.

The "secondary shifts" of ^{13}Cα, ^{13}Cβ, and ^{13}CO carbons from their residue-specific random coil values have been successfully used to determine the secondary structure of many proteins via the CSI method (Wishart & Sykes, 1994). In order to predict secondary structure on the basis of the chemical shifts obtained for deuterated, methyl-protonated proteins produced using the pyruvate scheme discussed previously (Rosen et al., 1996), it is necessary to modify the random coil values of each site in each residue with the values listed in Table 1. Using the Cα secondary shift data obtained from the chemical shifts reported for ∼70% deuterated Trp repressor bound to DNA (Yamazaki et al., 1994a) and the modified random coil values that include the effects of deuteration, we were able to accurately identify all six α-helices in TrpR, although the length of each of these was underestimated by approximately one residue at both the N- and C-terminal ends relative to a crystal structure of the TrpR/DNA complex (Lawson & Carey, 1993). Additionally, a small section of β-strand was incorrectly predicted in the immediate vicinity of the 5-methyltryptophan cofactor; this aromatic group likely leads to aberrant upfield shifts of the ^{13}Cα resonances, resulting in this misidentification. Note that the Cα chemical shifts reported for this complex are for C$\alpha-^2$H pairs (not C$\alpha-^1$H), and therefore, the corrections listed in Table 1 are, at least to first order, correct. However, because this sample was prepared using glucose as the carbon source and 70% D_2O, the levels of deuteration at the Cβ position and for carbons proximal to Cβ are likely to be different from that of the pyruvate-derived protein. The correction factors for the Cβ positions listed in Table 1 are not applicable in this case, and we have therefore not included the Cβ shifts in this analysis.

The random coil data base with the ^{13}Cα and ^{13}Cβ modifications indicated in Table 1 was used to calculate the secondary structure of the PLCC SH2 from the chemical shifts obtained on the deuterated, methyl-protonated sample. The CSI-based secondary structure was consistent with that observed in the high-precision reference structures of the PLCC SH2 (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation), although the lengths of most elements were underestimated by one or two residues at each end. A notable exception to this was the αB-helix, which was significantly truncated from 10 residues in the reference structures to only the four N-terminal residues by the CSI method. A CSI analysis of a fully protonated PLCC SH2 similarly underestimates the length of this helix, and of interest is the fact that the $^3J_{HN\alpha}$ coupling constants from this sample are all greater than 7 Hz in the C-terminal four residues of this helix (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation). It is also noteworthy that CSI predictions of secondary structural elements in the Grb2 SH2 domain have also severely underestimated the length of the αB helix (Wang et al., 1996).

Effect of Deuteration on Spectral Resolution. In the analysis of larger protein systems by NMR, adequate spectral resolution is a vital requirement. As observed in our previous work, the combined effect of pyruvate transaminases and the intrinsic ability of the pyruvate methyl protons to exchange with solvent deuterons leads to incomplete protonation of the target methyl groups, generating a mixture of CH_3, CH_2D, CHD_2, and CD_3 isotopomers at each of these sites (Rosen et al., 1996). Each methyl deuteron leads to upfield shifts of both the carbon nucleus (−0.3 ppm per ^{2}H) and the remaining methyl protons (−0.02 ppm per ^{2}H), giving rise to as many as three peaks per methyl group in carbon−proton correlation spectra (CD_3 is not detected in these experiments). Because spectral overlap becomes more critical with larger proteins, it is clearly important to minimize the effects of the methyl isotopomers as much as possible.

Many of the multidimensional experiments used for assignment of chemical shifts and NOE distances are limited in resolution in the indirectly detected dimensions. This is particularly the case in the carbon dimensions of 4D NOE-based experiments, where acquisition times of 4−5 ms are typically employed. The resolution in the carbon dimensions of these experiments is limited to a large extent, therefore, by the short acquisition times and not by the isotopomer distribution discussed above. To illustrate this point, we have recorded constant time ^{13}C−^{1}H HSQC spectra on both fully protonated and deuterated, methyl-protonated samples of the PLCC SH2. Although these spectra were obtained with acquisition times of $\sim 1/J_{CC}$ in the carbon dimension, where J_{CC} is the one-bond aliphatic carbon−carbon coupling constant, we have truncated the data in the carbon dimension after 4.4 ms of evolution during the processing stage. In this way, the resolution in the carbon dimensions of each data set is the same as, or in fact slightly better than, that obtained in the final processed 4D experiments. Note that in the 4D experiments evolution of magnetization due to one-bond homonuclear carbon couplings proceeds during the carbon evolution time. Panels a and b of Figure 2 illustrate that for the methyl groups arising from Ala, Val, Leu, and Ile (γ2) of the deuterated, methyl-protonated sample there is a small but noticeable degradation in spectral resolution relative to that of the corresponding spectrum of the fully protonated molecule. It is clear that, to a large extent, the poor resolution in the carbon dimension minimizes the deuterium isotope effects associated with the isotopomer distribution. In addition, however, NMR and mass spectroscopic data indicate that the average level of labeling of the pyruvate-derived amino acids Val, Ile (γ2), and Leu is on the order of two protons per methyl, and approximately 50% of these protons are derived from the CH_3 isotopomer (Rosen et al., 1996). Thus, the centers of the isotopomer distribution are heavily weighted toward the CH_3 methyl resonance position. In contrast, the Ile δ1 CH_3 groups, which are

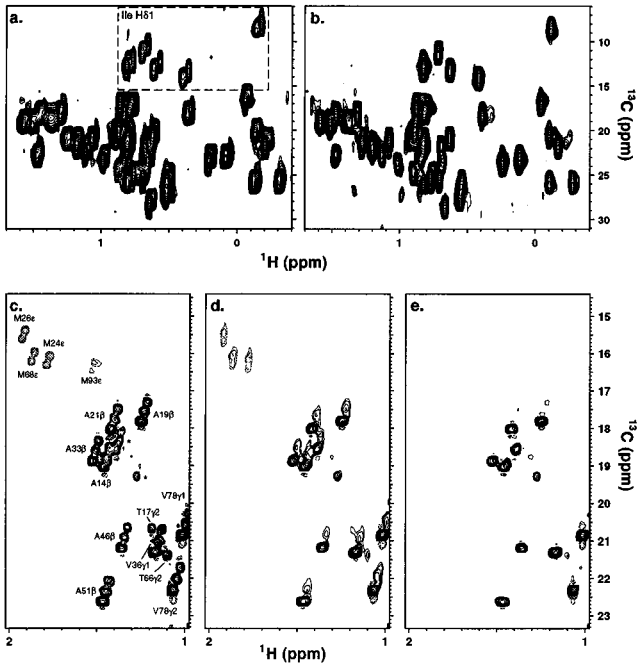

FIGURE 2: Effects of the deuterium isotopomer distribution on spectral resolution. (a and b) Methyl region of a constant time ^{13}C−^{1}H HSQC experiment recorded on (a) an ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl)-labeled PLCC SH2/pY1021 complex and (b) an ^{15}N, ^{13}C, ^{1}H-labeled version of the same complex. Spectra are derived from only the initial 4.4 ms of the carbon evolution [a total carbon evolution time of 28 ms was recorded for each spectrum, with ^{2}H decoupling in the case of the ^{15}N,^{13}C, ^{2}H, (^{1}H-methyl) PLCC SH2 sample]. The Ile Cδ1 region is identified by a dashed box. Spectra are plotted at similar distances from the noise floor. (c−e) Resolution improvement by eliminating cross-peaks from deuterium-containing methyl groups. All panels show a portion of the methyl region of constant time ^{13}C−^{1}H HSQC experiments recorded on a sample of the ^{15}N, ^{13}C, ^{2}H (^{1}H-methyl)-labeled PLCC SH2/pY1021 complex. These experiments utilized (c) ^{2}H broad-band decoupling during the ^{13}C evolution period ($1/J_{CC} \sim 28$ ms) for maximum resolution of the isotope-shifted peaks, (d) no ^{2}H broad-band decoupling to selectively broaden the CH_2D and CHD_2-derived cross-peaks and (e) a purging scheme described in the text to further reduce the intensity of CH_2D and CHD_2-derived cross-peaks. Spectra c−e are plotted at the same levels and were recorded at 25 °C.

derived from Thr and not pyruvate, are composed of approximately equal amounts of CH_3 and CHD_2 isotopomers (CH_2D is absent), and in this case, the deuterium isotope shift is much more pronounced. Note that cross-peaks from methyl groups that are not directly derived from pyruvate are at a much lower intensity (5−15-fold lower than those of Val and Leu) than peaks from pyruvate-derived residues.

Resolution is improved in ^{13}C−^{1}H correlation experiments by recording the carbon chemical shift in a constant time manner, where the duration of the constant time period is chosen to minimize sensitivity losses arising from carbon−carbon scalar couplings (Santoro & King, 1992; Vuister & Bax, 1992). Thus, maximum acquisition times in the carbon dimension are typically chosen to be multiples of the inverse of the one-bond aliphatic carbon−carbon coupling. In this case, the resolution is sufficient to allow observation of each of the methyl isotopomers, as illustrated in Figure 2c recorded with ^{2}H decoupling during the evolution of carbon magnetization. A simple way of minimizing the intensity from CH_2D and CHD_2 isotopomers is to eliminate ^{2}H decoupling, as indicated in Figure 2d. In this case, deuterium spin flips mix carbon multiplet components associated with different deuteron spin states, leading to a significant broadening and hence attenuation of these cross-peaks (Grzesiek et al., 1993; Grzesiek & Bax, 1994; London et al., 1994). Note that neither the Met Hϵ nor the Thr Hγ methyl groups are derived from pyruvate; as a consequence, neither has a CH_3 methyl peak and all cross-peaks from these positions are severely broadened. To further reduce the intensity of any remaining peaks from deuterium-containing methyl groups, it is possible to actively purge magnetization from these methyls. In this case, we have employed a strategy which is similar in some respects to that discussed previously in the context of selecting for CH_2D methyl groups (Muhandiram et al., 1995). Magnetization transferred from methyl protons to carbon at the start of the constant time carbon evolution period is allowed to evolve for a period of $1/(2J_{CH})$, where J_{CH} is the one-bond ^{13}C−^{1}H coupling. At this point, magnetization associated with $^{13}CH_3$, $^{13}CH_2D$, and $^{13}CHD_2$ spin systems is of the form $C_{TR}I_z^iI_z^j$ ($i \neq j$), $C_{TR}I_z^i$, and C_{TR}, respectively, where C_{TR} is transverse carbon magnetization and I_z^i refers to the z component of magnetization associated with methyl proton i. Application of a proton 90_x90_ϕ pulse pair at this point, where the phase of ϕ is incremented by 180° in successive scans with no change in the receiver phase, effectively eliminates magnetization from the $^{13}CH_2D$ group. Additional purging of signals arising from partially deuterated isotopomers is achieved by applying a ^{2}H purge (90°) pulse at a time of $1/(4J_{CD})$ after the start of the constant time carbon evolution period. In a previous publication, we have described how each of the carbon multiplet components in CD and CD_2 spin systems evolves due to carbon−deuterium scalar coupling(s) (Muhandiram et al., 1995). Neglecting deuterium spin flips for the moment, only the outer components of the C−D triplet are effected by J_{CD}, and they evolve at a rate of $\omega = \pm 2\pi J_{CD}$. In the case of the CD_2 spin system, evolution is more complex since the carbon multiplet is a 1:2:3:2:1 pentet. As before, the central line does not evolve, while the lines closest to and farthest from the central component evolve with angular frequencies of $\pm 2\pi J_{CD}$ and $\pm 4\pi J_{CD}$, respectively. It is therefore not possible to eliminate both sets of lines with a single purging pulse. Moreover, evolution is complicated by finite deuterium lifetimes which result in an interchange of magnetization between each of the multiplet components. In Figure 2e, we have chosen a compromise delay of $1/(4J_{CD})$ in the purging scheme employed, and the efficiency of purging suggests that this choice is reasonable.

Although the purging scheme described above could be incorporated into NOESY-based experiments, the use of constant time periods and, more importantly, the decrease in intensity of methyl NOEs arising from the purging of CH_2D and CHD_2 groups lead to unacceptable sensitivity losses. As such, the optimal approach for simplifying the distribution of methyl cross-peaks is to improve the yield of fully protonated methyl groups at the point of protein biosynthesis. Deuterium incorporation into the methyl groups appears to proceed by two mechanisms, one of which is catalyzed by pyruvate transaminase (Rosen et al., 1996). We have found that inducing protein overexpression for a shorter time at a higher temperature (9 h at 37 °C vs 24 h at 30 °C used in the present study) generated a PLCC SH2 sample with higher amounts of fully protonated methyl groups. Preliminary results demonstrate a 30% increase in the CH_3:CH_2D ratio of several pyruvate-derived methyl groups, suggesting that this shorter induction period ensures that the overexpressed protein incorporates pyruvate-derived amino acids before the internal pyruvate stores become

extensively deuterated. Other options currently under exploration include the use of transaminase-deficient bacterial strains, growth of bacterial cultures in the presence of transaminase inhibitors, and the development of an approach based on organic synthesis and enzymology to produce large quantities of ^{15}N, ^{13}C, ^{2}H-labeled methyl-protonated amino acids.

PLCC SH2 NOESY Analysis and Structure. In order to evaluate the strategy of obtaining global folds based on a limited data set of NH−NH, methyl−NH, and methyl−methyl distances, we have examined the precision and accuracy of structures obtained from experimental distance and dihedral angle restraint information derived from a deuterated, methyl-protonated sample of the PLCC SH2. Interproton distance restraints were obtained from three isotope-edited 4D NOESY experiments recorded with mixing times of 175 ms at 30 °C. The absence of large numbers of protons in highly deuterated systems dramatically limits the number of observable NOEs. As such, each individual interproton distance restraint has a significant influence on the total group of structural restraints. With this in mind, we have used 4D isotope-edited NOESY experiments so that two chemical shifts could be recorded to assist in the assignment of each proton of an interacting pair and have required unambiguous assignment of both diagonally related cross-peaks for a given NOE in the ^{15}N, ^{15}N- and ^{13}C, ^{13}C-edited NOESY experiments.

It is instructive to compare the pattern of observed NOEs to the distances present in the mean PLCC SH2 structure calculated from the highly refined reference structures (Figure 1). A high percentage of the possible NH−NH NOE cross-peaks were observed at distances up to 5 Å, but few were seen past this range. Previous studies using proteins deuterated uniformly and to high levels have noted NOEs between protons separated by longer distances, facilitated by the increased relaxation time of diagonal magnetization and the reduction in the number of spin-diffusion pathways (LeMaster & Richards, 1988; Venters et al., 1995). Structure calculations performed by both our group (discussed below) and others (Venters et al., 1995) suggest that it is not feasible to obtain accurate global folds using only distance restraints between backbone NH protons of 5 Å or less. As such, more distance information is needed; with deuterated, methyl-protonated systems, additional distances are available from methyl−NH and methyl−methyl NOEs. As shown in Figure 1, a significant percentage of the possible cross-peaks for these NOEs were observed for distances greater than 5 Å, and this is especially important given the relatively large number of possible restraints available at these distances.

To demonstrate the improvements gained by incorporating methyl−methyl and methyl−NH NOE information, two sets of structures differing in the types of restraint information used were generated. One group, NN, used distance restraints only between backbone NH protons combined with backbone dihedral restraints obtained from the CSI analysis of $^{13}C\alpha$, $^{13}C\beta$, and ^{13}CO chemical shifts; another, CN, also utilized the additional methyl−NH and methyl−methyl NOEs available from our labeling scheme. The combined numbers of distance and dihedral angle restraints per residue in the NN and CN groups are approximately two and four, respectively. Note that these restraint densities are significantly below those of first-generation NMR structures (Clore & Gronenborn, 1991). In contrast, the reference set of PLCC SH2 structures were generated on the basis of approximately 18 restraints per residue.

The average number of restraints per residue is an overly simplistic descriptor of the available distance restraints, however, since the influence of NOEs depends not only on their number but also on their placement within the structure (Liu et al., 1992). In this case, NOEs involving the protonated methyl groups are quite important due to the location of many methyl groups in protein hydrophobic cores, where they are near other protons that are distant in the primary structure. This is emphasized in a schematic representation of the PLCC SH2 with the protonated methyl groups indicated using a space-filling representation (Figure 3a). The majority of methyl groups are located between the central β-sheet and two flanking α-helices, placing the methyl protons of one secondary structure element near protons from other elements. This is reflected in the NOEs experimentally observed from the deuterated, methyl-protonated PLCC SH2 sample; the median NH−NH NOE is between protons on sequential residues, while the median methyl−NH NOE involves protons separated by two residues. In contrast, the median methyl−methyl NOE is separated by 14 residues (for all CH_3-CH_3 NOEs; 17 residues if only the interresidue methyl NOEs are considered). These NOEs are thus extremely valuable even when the distance between the methyls is imprecisely known. Correspondingly, a much greater percentage of the methyl−methyl NOEs are long-range compared to the other two classes of NOEs ($|i - j| \geq 5$ residues): NH−NH (18%), methyl−NH (36%), and methyl−methyl (72% of all NOEs, 88% of inter-residue NOEs).

Examination of structures determined (i) on the basis of NH−NH NOEs only, (ii) from NH−NH, methyl−NH, and methyl−methyl NOEs, and (iii) from a full complement of NOEs derived from a fully protonated sample illustrates the importance of the methyl-based distance restraints. The NN class of structures (Figure 3b) have a fairly well-formed central β-sheet (rmsd to the reference mean, 3.38 ± 0.80 Å), due to the presence of many interstrand NH−NH NOEs. Similar accuracy is observed within either of the α-helices, restrained by (i, i+1) and (i, i+2) NH−NH NOEs. However, the distance between backbone NH atoms from different secondary structural elements is usually greater than the 5 Å maximum distance that was experimentally observed in the ^{15}N, ^{15}N-edited NOESY data set. As such, few NOEs have been obtained between the helices and the central sheet, effectively leaving them disordered with respect to each other. These results are consistent with those previously found for human carbonic anhydrase and the HIV-1 Nef protein (Grzesiek et al., 1995; Venters et al., 1995). The locations of the protonated methyl groups in positions between the central sheet and flanking helices (Figure 3a) result in significant numbers of NOEs between protons from residues in different secondary structural elements. This is observed in the CN group of structures that were generated by including methyl−NH and methyl−methyl distance restraints (Figure 3c). Note that the helices are well-oriented with respect to the central sheet and the overall shape of the SH2 domain much more closely resembles that of the reference structures (Figure 3d).

The improved accuracy of structures utilizing NOEs involving methyl groups is dramatically illustrated in Figure 4 where the average reference structure is superimposed with the average structure determined from NN restraints (a) or

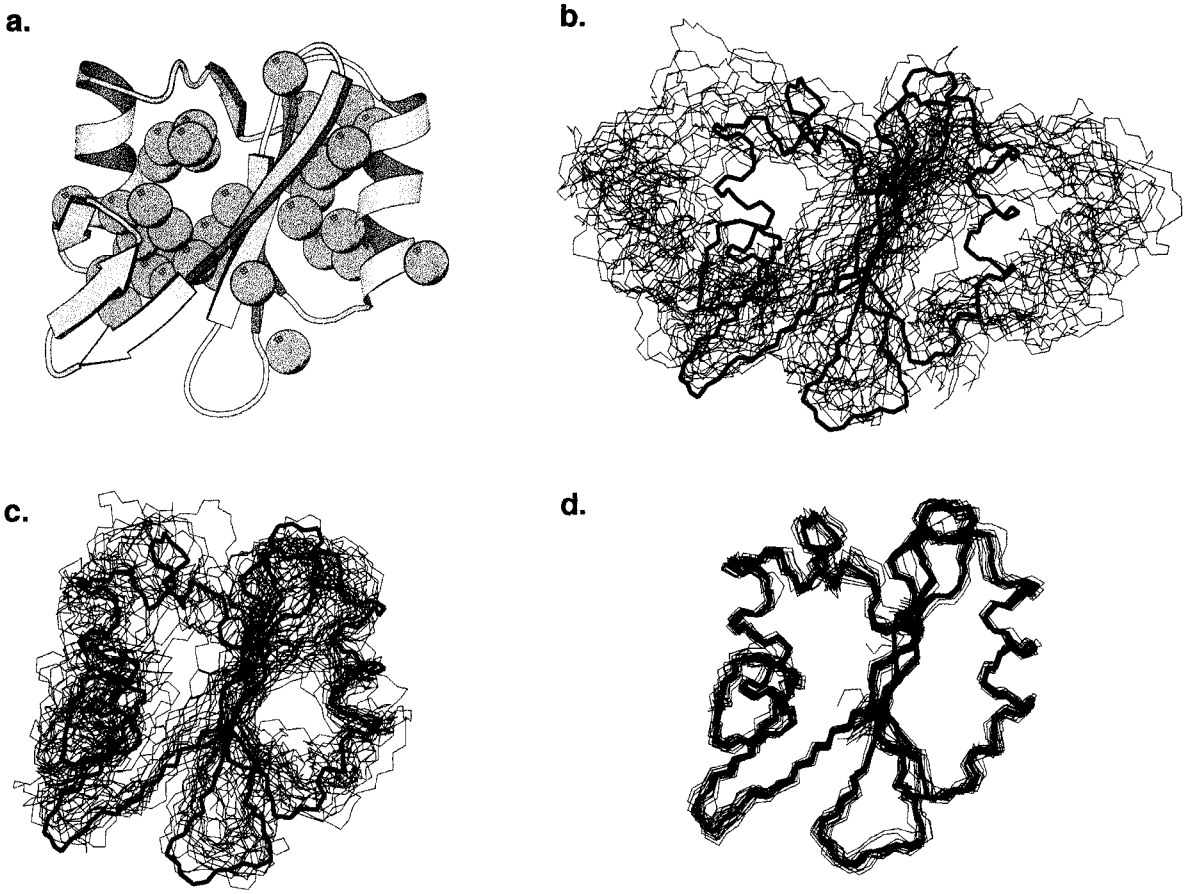

FIGURE 3: Structures of the PLCC SH2. (a) Schematic view showing ordered secondary structure elements identified by Pascal et al. (1994). Pyruvate-derived methyl groups are shown in a space-filling representation, highlighting their position between the central β-sheet and flanking α-helices. (b) NN group of structures [1.1 distance and 0.7 dihedral restraint(s)/residue]. (c) CN group of structures [2.7 distance and 0.7 dihedral restraint(s)/residue]. (d) High-precision reference structures, generated using an average of 16.0 distance and 1.7 dihedral angle restraints per residue (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation). Each bundle has 18 independently refined structures (thin lines) and is fit to the mean reference structure (thick line) using the backbone atoms of the secondary structure elements shown in panel a. Panel a is drawn using MOLSCRIPT (Kraulis, 1991); all other figures containing molecular graphics were generated using MOLMOL (Koradi et al., 1996).

the average structure generated from the CN set of restraints (b). Note that in both cases (NN or CN set of structures) the (ϕ,ψ) dihedral restraints were included in the manner described above. As already discussed, the NN group of structures has a fairly well-defined central β-sheet but the poor definition of the helix/sheet interface results in helices which are distorted beyond recognition in the average structure (Figure 4a). In contrast, the average structure from the CN set of restraints superimposes quite well on the average reference structure, with an rmsd value of 2.03 Å for the backbone atoms in all secondary structure elements as defined in the Table 2 legend (Figure 4b). The primary differences between the two structures depicted in Figure 4b are localized to loop regions connecting secondary structure elements.

The improvements in precision and accuracy of structures generated with the inclusion of methyl NOEs are quantitated in Table 2. The precision of PLCC SH2 structures is significantly improved upon the addition of the methyl−NH and methyl−methyl distance restraints; the rmsd among the CN group members is approximately 1.5 times smaller than among NN members. As seen in Figures 3 and 4, the ordered secondary structure elements are far more easily recognized as a result of this improvement. More important, though, is the dramatically improved accuracy. A comparison of the structures from the NN and CN groups shows an approximate 2.5-fold improvement in the backbone rmsd with respect to the PLCC SH2 reference structure. The overall shape of the molecule is much more accurately determined as well, as shown by the better agreement between the radius of gyration of the mean PLCC SH2 structure ($R_g = 12.2$ Å) and that of the CN group of structures (12.3 ± 0.2 Å) compared to those of the structures generated from NH−NH restraints only (15.5 ± 0.5 Å). The primary goal of a structure determination should be to produce an accurate set of structures, as opposed to a precise set. Inaccuracies lead to incorrect conclusions that may not be realized until correct data are available, while imprecision should somewhat more innocuously prevent conclusions from being reached in the first place. This principle is emphasized by the utility of structures and models produced from inherently lower-resolution structural methods (e.g. electron diffraction) that may not have the precision of crystallographically or NMR-determined structures but can still provide useful insight into biological function if they are accurate.

As noted above, the NOE-based distance restraints have been supplemented in these structure calculations with restraints on the ϕ and ψ backbone dihedral angles obtained by CSI-based identification of secondary structure elements. The influence of dihedral angle restraints on the precision and accuracy of these structures was determined by calculating a set of PLCC SH2 structures using the CN group of NOE restraints without any dihedral angle restraints. Omit-

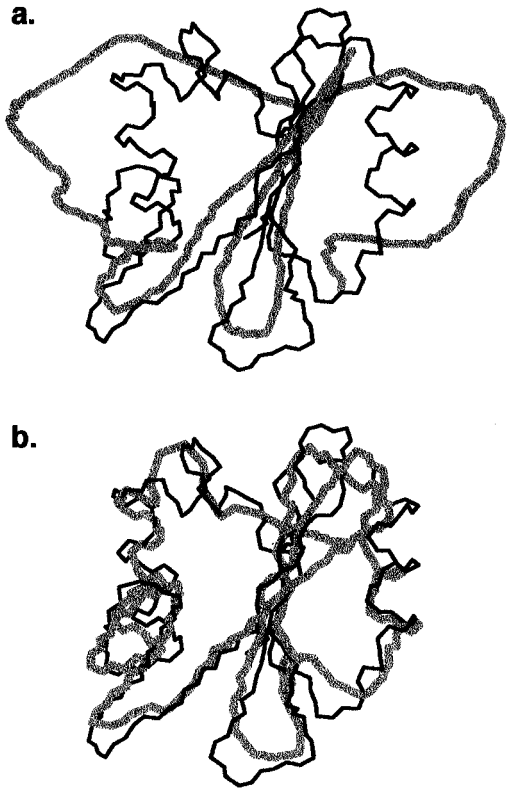

FIGURE 4: Accuracy of PLCC SH2 average structures determined using either the NN or CN distance restraint sets. (a) Overlay of the average structure of the 18 NN group structures (thick gray line) with the average of the 18 high-precision reference structures (thin black line) (S. M. Pascal, A. U. Singer, L. E. Kay & J. D. Forman-Kay, manuscript in preparation). A 6.91 Å rmsd between the average structures of the NN and high-precision reference sets of data is calculated considering the backbone atoms of residues in secondary structure elements (as defined in the Table 2 legend). (b) Overlay of the average structure of the 18 CN group structures (thick gray line) with the average high-precision reference structure (thin black line). The rmsd of backbone atoms in secondary structure elements is 2.03 Å. Structures were aligned by minimizing the rmsd between backbone atoms in residues 10−100 and are shown from the same perspective as Figure 3.

ting the dihedral restraints leads to a slight decrease in the precision of residues identified by CSI to be in secondary structural elements (Table 2), although the precision of residues 10−100 in the domain was not affected. However, the accuracy was more strongly dependent on the inclusion of the CSI-derived ϕ and ψ restraints, as indicated in Table 2.

Structure Determinations of Larger Proteins Based on Simulated Restraint Data Sets. To determine the potential of NH−NH, methyl−NH, and methyl−methyl NOE restraints in determining global folds of larger proteins with different secondary structure compositions, we generated structures of six other proteins from simulated distance and dihedral angle restraint sets (Table 3). The proteins are all monomeric and display an array of different structural topologies. The restraint tables were created on the basis of the amount and precision of data experimentally obtained from the studies of the PLCC SH2. The distance restraints used were generated on the basis of the distance- and type-dependent (i.e. NH−NH, methyl−NH, and methyl−methyl) frequency of observation of the particular class of NOE seen in experimental data sets recorded on the ^{15}N, ^{13}C, ^{2}H (methyl-protonated) PLCC SH2 (Figure 1), with five different sets of distance restraints randomly generated for each protein. Note that this approach effectively considers a best case scenario where the increasing peak overlap of larger systems does not affect the percentage of assignable NOEs. We are currently developing further biosynthetic strategies coupled with NMR methodology to address problems of assignment of NOEs arising from overlap in spectra of large proteins.

As discussed above, the low density of restraints (approximately two to four restraints/residue) leads to structures which are highly influenced by the presence or absence of small numbers of long-range restraints. With the inclusion of structures from these different restraint sets, the quality of the structures determined by this method can be evaluated without unfairly biasing the results toward or away from any one set of randomly selected restraints. Backbone dihedral angle restraints were included in the calculations by using all residues identified in each crystal structure as residing in an α-helix or β-sheet conformation and assigning ϕ and ψ ranges to each dihedral angle as discussed above. In an experimental NMR structure determination of a highly deuterated protein, this information can be derived from the ^{13}C CSI approach, as indicated above for the PLCC SH2. A recent statistical evaluation of the efficiency of ^{13}C CSI-based secondary structure identification (Luginbühl et al., 1995) suggests that this method will only identify approximately 50% of β-strand and 75% of α-helical residues; a variety of methods could be used to complement this approach to obtain a more complete assignment, including the analysis of sequential NH−NH NOE intensities, amide ^{1}H−^{2}H exchange rates, and scalar coupling constants.

As anticipated from the PLCC SH2 case, the incorporation of additional distance information in these structure determinations significantly improves the precision and accuracy of the resulting structures. The methyl−NH and methyl−methyl distance information provides a 1.4-fold increase in the number of distance restraints available over the NH−NH list alone and more significantly often provides connectivities between residues that are well separated in primary structure. With the inclusion of this additional information, the accuracy and precision of the CN class of structures are both at least a factor of 2.5-fold improved over the corresponding NN class for each of the proteins studied. It is noteworthy that the radii of gyration (R_g) of the CN class of structures and the corresponding crystal structures used in each determination are within error while the R_g values of the NN structures are up to 40% larger (data not shown). This general trend of under-restrained systems having expanded volumes appears to be a product of the structure determination process when an insufficient number of long-range restraints are used, suggesting that monitoring the degree of packing of structures during a de novo determination may offer an independent assessment of the accuracy of these structures.

On the basis of the six proteins examined, it is apparent that both secondary structure content and topology play significant roles in determining the quality of the structures that can be obtained from the limited class of NOEs that we are considering. This is especially evident with the two highly helical systems (1gmf and 1fps); structures in the NN class, lacking the longer-range methyl-based distance restraints, are comprised of highly extended sets of helices. This is likely due to the long distances between NH protons on adjacent helices (minimum separation of approximately 6 Å) and the fact that helices often cross each other at fairly sharp angles, further reducing the pairs of NH protons that

Table 2: Precision and Accuracy of PLCC SH2 Structures Generated with Different Amounts of Restraint Information[a]

	precision (Å)			accuracy (Å)			R_g (Å)
	residues 10−100	2° (all)	2° (CSI)	residues 10−100	2° (all)	2° (CSI)	residues 10−100
reference (1874/197)[b]	0.66 ± 0.11	0.48 ± 0.08	0.41 ± 0.07	na	na	na	12.2 ± 0.1
CN (287/72)	2.87 ± 0.45	2.36 ± 0.45	1.88 ± 0.44	4.13 ± 0.71	3.08 ± 0.40	2.50 ± 0.35	12.3 ± 0.2
CN, no dihedrals (287/0)	2.87 ± 0.48	2.52 ± 0.45	2.18 ± 0.47	4.44 ± 0.56	3.65 ± 0.43	3.26 ± 0.51	12.3 ± 0.2
NN (112/72)	3.78 ± 0.52	3.53 ± 0.50	3.32 ± 0.55	8.32 ± 0.69	7.74 ± 0.75	6.77 ± 1.01	15.5 ± 0.5

[a] All statistics were generated by rms fitting of backbone (N, C, and Cα) atoms of specified regions of 18 independently generated structures to the appropriate target (precision, mean structure of 18; accuracy, mean structure of reference group). The CN and NN sets of structures had an average of 1.0 NOE restraint violated by 0.5 Å and 0.2 experimental dihedral restraint violated by 5°. Regions outside of residues 10−100 did not adopt any stable structure in previous studies (Pascal et al., 1994). Secondary structure regions are defined as the following: α-helical and β-strand regions [all, residues 11−13, 17−28, 34−38, 45−51, 54−63, 66−69, 72−74, 77−87, 97, and 98; from Pascal et al. (1994)] (CSI, residues 18−25, 34−37, 45−50, 54−63, 66−69, and 77−80; from this work). The accuracy statistics are not given for the reference structures as the accuracy is defined relative to the mean reference structure. The radius of gyration (R_g) was calculated for the Cα atoms of residues 10−100 using X-PLOR (Brünger, 1992). [b] Each group of structures is identified by the types of restraints being used (number of distance restraints/number of dihedral restraints).

Table 3: Precision and Accuracy of Structures Determined from Simulated Restraints[a]

structure	class	% α/β	molecular mass (kDa)	precision (Å)	accuracy (Å)
1gmf	α+β	50/7	13.6	**3.92** ± **1.10**/9.11 ± 2.23	**4.46** ± **1.53**/13.73 ± 1.91
1stn	α+β	23/27	15.5	**2.16** ± **0.31**/6.10 ± 1.73	**2.56** ± **0.34**/9.23 ± 3.44
1xnb	β	5/58	20.4	**2.91** ± **0.36**/4.70 ± 0.52	**3.64** ± **0.38**/7.50 ± 1.47
1mua	α+β	13/28	28.7	**3.54** ± **0.75**/8.67 ± 1.66	**5.11** ± **1.28**/13.00 ± 2.44
1fps	α	75/1	40.0	**6.18** ± **1.42**/24.96 ± 6.66	**7.54** ± **2.81**/>40.0
1mpb	α/β	45/16	40.7	**2.87** ± **0.65**/14.91 ± 1.90	**3.60** ± **0.59**/18.47 ± 1.83

[a] Structures are listed by their PDB reference code, topology class, percent content of α-helix and β-strand, and molecular mass as tabulated by MOOSE. The topology class is one of the following: all-helix (α), all-sheet (β), intermixed helix and sheet (α/β), or separated helix and sheet (α+β). Precision and accuracy are calculated for the backbone N, C, and Cα atoms from all secondary structure elements; this number can be significantly improved if individual subdomains are considered separately (example, for 1fps, precision of two subdomains independently is 3.18 ± 0.41 Å and 5.37 ± 1.68 Å for the CN group) or by the omission of a poorly defined secondary element near a terminus (example, for 1mua, dropping two short N-terminal helices changes precision to 2.20 ± 0.32 Å and accuracy to 2.72 ± 0.34 Å for the CN group). Precision and accuracy values in bold text are from CN data sets, and values in plain text are from NN sets.

are in spatial proximity. In this case, the approximate cutoff of 5 Å that has been observed for NH−NH NOEs in the PLCC SH2 and is used in the present analysis is too short to allow the observation of a sufficient number of NOEs to orient adjacent helices with respect to each other. These problems are not significant for the case of β-sheets, where interstrand NH−NH distances are on the order of 3 Å and are easily observed. As such, structures of mixed α/β and all-β proteins can be determined to higher accuracy than is the case for all-helical proteins.

Previous studies have reported that the accuracy and precision of protein structures generated by NOE distance restraints will be dependent on molecular mass, with larger proteins demonstrating worse statistics (Liu et al., 1992). This is based on the observation that as a function of increasing protein size the total number of interproton distances grows roughly as the square of the number of residues (N^2) while the number of short-range (≤6 Å) restraints increases only linearly with the number of residues (N). As such, the shorter-range distances represent a progressively smaller portion of the total distances as the molecular mass increases, leading to lower precision and accuracy. Our studies indicate however that the major determinant in structural precision and accuracy is a combination of both the relative amounts of ordered secondary structure elements and the topology that they adopt. This is demonstrated by differences in the quality of structures obtained from systems of comparable size (e.g. 1gmf/1stn and 1fps/1mpb) and the similarities in quality between systems of different molecular masses (e.g. 1xnb/1mpb).

An additional benefit from the incorporation of distance restraints to methyl protons is the significant ordering of the side chains of these methyl-containing residues. An example of this is shown for staphylococcal nuclease (Figure 5). Several hydrophobic clusters of these residues can be identified, with the Ala, Val, Ile, and Leu side chains having

FIGURE 5: Positions of methyl-protonated side chains within staphylococcal nuclease. The backbone of staphylococcal nuclease derived from the crystal structure (1stn.pdb) is drawn with thick lines (gray backbone and black Ala/Ile/Val/Leu side chains). Overlaid on top of this are the Ala/Ile/Val/Leu side chains (thin black lines) of 18 structures generated from simulated distance and dihedral angle restraints. Structures were oriented by aligning the backbone heavy atoms of each structure generated from simulated data with the analogous atoms in the reference crystal structure (rmsd = 2.56 ± 0.34 Å over all α-helical and β-strand residues). A total of 35 Ala/Ile/Val/Leu side chains are shown; an rmsd of 2.33 ± 0.95 Å between the carbon atoms of these side chains derived from the X-ray reference and the NMR-based structures is obtained.

accuracy comparable to that of the backbone (side chain heavy atom rmsd of 2.33 Å vs a backbone rmsd of 2.56 Å). The mean χ_1 accuracy (46°) of these residues also shows that these side chains are in the correct rotameric state. This level of determination of side chain conformation may facilitate efforts to model the locations of more poorly defined side chains into the core of these proteins.

Future Prospects. As demonstrated above, the additional structural information gained by the incorporation of distance restraints into protonated methyl groups significantly improves the quality of structures generated from experimental or simulated data sets. These additional restraints are obtained without compromising any of the experiments used for backbone chemical shift assignment. In the case of the PLCC SH2, 37 methyl sites are protonated to complement the 99 backbone NH sites.

A number of alternative strategies have recently been proposed for the incorporation of protons into highly deuterated proteins. Protonated (and potentially ^{15}N, ^{13}C-labeled) amino acids, such as Ile, Val, and Leu, can be added to bacteria grown on minimal D_2O-based media and deuterated glucose (Crespi et al., 1968; Brodin et al., 1989; Oda et al., 1992; Metzler et al., 1996; Smith et al., 1996). In vivo exchange processes replace the Hα proton with a deuteron, while the remainder of the side chain remains highly protonated. As such, this approach avoids fractional deuteration and the accompanying problems with distributions of chemical shifts from deuterium isotope effects. Additionally, one can produce samples from different combinations of protonated amino acids; judicious choices of combinations could help circumvent spectral overlap problems. Unfortunately, there are several drawbacks to this approach of producing deuterated, residue-protonated proteins. The high level of side chain protonation will complicate sequence-specific assignment of methyl groups (especially in high-molecular mass proteins) by reducing the efficiency of methods which utilize scalar-coupled magnetization transfer steps through ^{13}C. For large proteins, with efficient relaxation, assignment of side chain chemical shifts will therefore be based to some extent on NOE experiments. The additional protons will also lead to more effective spin diffusion in NOESY experiments and will eliminate some of the improvements in the sensitivity of backbone NH−NH NOEs that perdeuteration provides (LeMaster & Richards, 1988; Torchia et al., 1988). As well, spin diffusion within a side chain will lower the resolution of measured NOEs, requiring the use of large pseudoatoms or similar corrections. Finally, the addition of intraresidue magnetization transfer pathways will lead to a decrease in the intensities of important long-range NOE contacts. In the ideal sample, these problems would be avoided by complete deuteration of all non-methyl side chain positions while completely protonating the selected methyl groups. An approach based on incorporation of synthetically or enzymatically prepared amino acids that are highly deuterated at all but the methyl positions may well prove to be the most practical. In any event, the conclusions of the present study regarding the utility of methyl NOEs for obtaining global folds are independent of the strategy used for the incorporation of such labeling.

Additional sources of information from deuterated, methyl-protonated samples are potentially available to further improve the quality of structures. To this point, we have investigated only a conservative set of the possible sources of distance and dihedral angle information. Several classes of information can improve the definition of secondary structure elements, including explicit hydrogen bond restraints and tighter bounds on backbone dihedral restraints established through the measurement of heteronuclear and homonuclear scalar couplings (Bax et al., 1994). Similar measurements could also be used to provide side chain angle restraints, potentially in conjunction with short-range methyl−NH NOEs. Several other sites in pyruvate-derived proteins are sufficiently protonated to obtain useful NOEs; these include nonexchangeable protons derived from the pyruvate methyl group via metabolic intermediates (Ser Hβ, Asn/Asp Hβ, Gln/Glu Hγ, Arg Hγ, and Pro Hγ) as well as exchangeable protons on side chain NH and NH_2 groups. A suite of NMR experiments has recently been developed to assign the chemical shifts of these exchangeable protons and demonstrated on a highly deuterated sample of the 29 kDa human carbonic anhydrase protein (Farmer & Venters, 1996). Finally, data base-oriented approaches that incorporate side chain preferences of conformation and environment (Kuszewski et al., 1996) could also improve the quality of structures determined from these samples.

In conclusion, we have used experimental and simulated distance and dihedral restraint sets to demonstrate the quality of structures that are available from NMR analyses of deuterated, methyl-protonated proteins. As a result of the lower numbers of interproton distance restraints available from these systems, the precision of structures is substantially lower than what can be obtained from studies of smaller ^{15}N, ^{13}C, ^{1}H-labeled proteins. However, the quality of structures generated is useful for a wide variety of applications, including topology recognition and guiding site-specific mutagenesis experiments. With heteronuclear-filtering techniques, the complexes of deuterated, methyl-protonated proteins with small ligands or other proteins can also be obtained, further extending the applicability of this work.

ACKNOWLEDGMENT

The authors thank Randall Willis (Hospital for Sick Children, Toronto) and Wendy Parris (Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto) for their assistance with sample preparation, Dr. Steven Pascal, Dr. Julie Forman-Kay, and Alex Singer (Hospital for Sick Children) for providing NMR and structural data on the refined PLCC SH2 structures prior to publication, and Dr. Gerry Gish (Samuel Lunenfeld Research Institute, Mount Sinai Hospital) for providing phosphorylated peptides used with the PLCC SH2 work. The authors are grateful for the enthusiastic support of Dr. Tony Pawson (Samuel Lunenfeld Research Institute, Mount Sinai Hospital) throughout this project. Evan Stein and Dr. Greg Warren (Yale University, New Haven, CT) are thanked for providing X-PLOR scripts and comments on the use of dihedral angle-based molecular dynamics.

REFERENCES

Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., & Zhu, G. (1994) *Methods Enzymol. 239*, 79−105.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Brodin, P., Drakenberg, T., Thulin, E., Forsén, S., & Grundström, T. (1989) *Protein Eng. 2*, 353−358.

Brünger, A. T. (1992) *X-PLOR 3.1 Manual*, Yale University Press, New Haven.
Campbell, R. L., Rose, D. R., Wakarchuk, W. W., To, R., Sung, W., & Yaguchi, M. (1993) in *Proceedings of the Second TRICEL Symposium of Trichoderma Reesi Cellulases and Other Hydrolases* (Suominen, P., & Reinikainen, T., Eds.) pp 63−72, Foundation for Biotechnical and Industrial Fermentation Research, Espoo.
Clore, G. M., & Gronenborn, A. M. (1991) *Science 252*, 1390−1399.
Clore, G. M., & Gronenborn, A. M. (1994) *Methods Enzymol. 239*, 349−363.
Crespi, H. L., & Katz, J. J. (1969) *Nature 224*, 560−562.
Crespi, H. L., Rosenberg, R. M., & Katz, J. J. (1968) *Science 161*, 795−796.
Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
Diederics, K., Boone, T., & Karplus, P. A. (1991) *Science 254*, 1779−1782.
Farmer, B. T., II, & Venters, R. A. (1995) *J. Am. Chem. Soc. 117*, 4187−4188.
Farmer, B. T., II, & Venters, R. A. (1996) *J. Biomol. NMR 7*, 59−71.
Farrow, N. A., Muhandiram, D. R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., & Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
Gardner, K. H., Konrat, R., Rosen, M. K., & Kay, L. E. (1996) *J. Biomol. NMR 8*, 351−356.
Grzesiek, S., & Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
Grzesiek, S., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
Grzesiek, S., & Bax, A. (1994) *J. Am. Chem. Soc. 116*, 10196−10201.
Grzesiek, S., Anglister, J., Ren, H., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 4369−4370.
Grzesiek, S., Wingfield, P., Stahl, S., Kaufman, J. D., & Bax, A. (1995) *J. Am. Chem. Soc. 117*, 9594−9595.
Hansen, P. E. (1988) *Prog. NMR Spectrosc. 20*, 207−255.
Hynes, T. R., & Fox, R. O. (1991) *Proteins: Struct., Funct., Genet. 10*, 92−105.
Janin, J., Miller, S., & Chothia, C. (1988) *J. Mol. Biol. 204*, 155−164.
Johnson, B. A., & Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−614.
Kay, L. E., Bull, T., Nicholson, L. K., Griesinger, C., Schwalbe, H., Bax, A., & Torchia, D. (1992a) *J. Magn. Reson. 100*, 538−558.
Kay, L. E., Keifer, P., & Saarinen, T. (1992b) *J. Am. Chem. Soc. 114*, 10663−10665.
Kay, L. E., Xu, G. Y., & Yamazaki, T. (1994) *J. Magn. Reson. A 109*, 129−133.
Koradi, R., Billeter, M., & Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kuszewski, J., Gronenborn, A. M., & Clore, G. M. (1996) *Protein Sci. 5*, 1067−1080.
Lawson, C. L., & Carey, J. (1993) *Nature 366*, 178−182.
LeMaster, D. M., & Richards, F. M. (1988) *Biochemistry 27*, 142−150.
Liu, Y., Zhao, D., Altman, R., & Jardetzky, O. (1992) *J. Biomol. NMR 2*, 373−388.
London, R. E., LeMaster, D. M., & Werbelow, L. G. (1994) *J. Am. Chem. Soc. 116*, 8400−8401.
Luginbühl, P., Szyperski, T., & Wüthrich, K. (1995) *J. Magn. Reson. B 109*, 229−233.
Markley, J. L., Putter, I., & Jardetzky, O. (1968) *Science 161*, 1249−1251.
Metzler, W. J., Wittekind, M., Goldfarb, V., Mueller, L., & Farmer, B. T., II (1996) *J. Am. Chem. Soc. 118*, 6800−6801.
Muchmore, S. W., Sattler, M., Liang, H., Meadows, R. P., Harlan, J. E., Yoon, H. S., Nettesheim, D., Chang, B. S., Thompson, C. B., Wong, S.-L., Ng, S.-C., & Fesik, S. W. (1996) *Nature 381*, 335−341.
Muhandiram, D. R., & Kay, L. E. (1994) *J. Magn. Reson. B 103*, 203−216.
Muhandiram, D. R., Xu, G. Y., & Kay, L. E. (1993) *J. Biomol. NMR 3*, 463−470.
Muhandiram, D. R., Yamazaki, T., Sykes, B. D., & Kay, L. E. (1995) *J. Am. Chem. Soc. 117*, 11536−11544.
Neri, D., Szyperski, T., Otting, G., Senn, H., & Wüthrich, K. (1988) *Biochemistry 28*, 7510−7516.
Nietlispach, D., Clowes, R. T., Broadhurst, R. W., Ito, Y., Keeler, J., Kelly, M., Ashurst, J., Oschkinat, H., Domaille, P. J., & Laue, E. D. (1996) *J. Am. Chem. Soc. 118*, 407−415.
Oda, Y., Nakamura, H., Yamazaki, T., Nagayama, K., Yoshida, M., Kanaya, S., & Ikehara, M. (1992) *J. Biomol. NMR 2*, 137−147.
Pascal, S. M., Singer, A. U., Gish, G., Yamazaki, T., Shoelson, S. E., Pawson, T., Kay, L. E., & Forman-Kay, J. D. (1994) *Cell 77*, 461−472.
Rice, L. M., & Brünger, A. T. (1994) *Proteins: Struct., Funct., Genet. 19*, 277−290.
Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T., & Kay, L. E. (1996) *J. Mol. Biol. 237*, 627−636.
Santoro, J., & King, G. C. (1992) *J. Magn. Reson. 97*, 202−207.
Schleucher, J., Sattler, M., & Griesinger, C. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 1489−1491.
Shan, X., Gardner, K. H., Muhandiram, D. R., Rao, N. S., Arrowsmith, C. H., & Kay, L. E. (1996) *J. Am. Chem. Soc. 118*, 6570−6579.
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., & Laue, E. D. (1996) *J. Biomol. NMR 8*, 360−368.
Stonehouse, J., Shaw, G. L., Keeler, J., & Laue, E. D. (1994) *J. Magn. Reson. A 107*, 178−184.
Tarshis, L. C., Yan, M., Poulter, C. D., & Sacchettini, J. C. (1994) *Biochemistry 33*, 10871−10877.
Torchia, D. A., Sparks, S. W., & Bax, A. (1988) *J. Am. Chem. Soc. 110*, 2320−2321.
Tsang, P., Wright, P. E., & Rance, M. (1990) *J. Am. Chem. Soc. 112*, 8183−8185.
Tweedy, N. B., Nair, S. K., Paterno, S. A., Fierke, C. A., & Christianson, D. W. (1993) *Biochemistry 32*, 10944−10949.
Venters, R. A., Metzler, W. J., Spicer, L. D., Mueller, L., & Farmer, B. T., II (1995) *J. Am. Chem. Soc. 117*, 9592−9593.
Vuister, G. W., & Bax, A. (1992) *J. Magn. Reson. 98*, 428−435.
Vuister, G. W., Clore, G. M., Gronenborn, A. M., Powers, R., Garrett, D. S., Tschudin, R., & Bax, A. (1993) *J. Magn. Reson. B 101*, 210−213.
Wang, Y. S., Frederick, A. F., Senior, M. M., Lyons, B. A., Black, S., Kirschmeier, P., Perkins, L. M., & Wilson, O. (1996) *J. Biomol. NMR 7*, 89−98.
Wishart, D. S., & Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., & Kay, L. E. (1994a) *J. Am. Chem. Soc. 116*, 11655−11666.
Yamazaki, T., Lee, W., Revington, M., Mattiello, D. L., Dahlquist, F. W., Arrowsmith, C. H., & Kay, L. E. (1994b) *J. Am. Chem. Soc. 116*, 6464−6465.
Yip, P. F. (1990) *J. Magn. Reson. 90*, 382−383.
Zhang, H., Zhao, D., Revington, M., Lee, W., Jia, X., Arrowsmith, C., & Jardetzky, O. (1994) *J. Mol. Biol. 238*, 592−614.
Zhou, M.-M., Ravichandran, K. S., Olejniczak, E. T., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., & Fesik, S. W. (1995) *Nature 378*, 584−592.

BI9624806