JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Quantifying Millisecond Exchange Dynamics in Proteins by CPMG Relaxation Dispersion NMR Using Side-Chain ^{1}H Probes

Alexandar L. Hansen,[†] Patrik Lundström,[‡] Algirdas Velyvis,[†] and Lewis E. Kay*[,†,§]

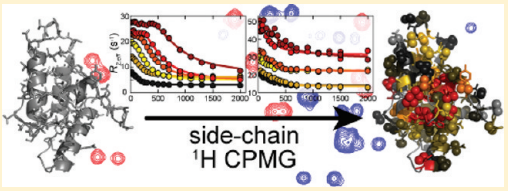
[†]Departments of Molecular Genetics, Biochemistry and Chemistry, The University of Toronto, Toronto, Ontario M5S 1A8, Canada

[‡]Division of Molecular Biotechnology, Department of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden

[§]Hospital for Sick Children, Program in Molecular Structure and Function, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada

Supporting Information

ABSTRACT: A Carr–Purcell–Meiboom–Gill relaxation dispersion experiment is presented for quantifying millisecond time-scale chemical exchange at side-chain ^{1}H positions in proteins. Such experiments are not possible in a fully protonated molecule because of magnetization evolution from homonuclear scalar couplings that interferes with the extraction of accurate transverse relaxation rates. It is shown, however, that by using a labeling strategy whereby proteins are produced using {^{13}C,^{1}H}-glucose and D_2O a significant number of 'isolated' side-chain ^{1}H spins are generated, eliminating such effects. It thus becomes possible to record ^{1}H dispersion profiles at the β positions of Asx, Cys, Ser, His, Phe, Tyr, and Trp as well as the γ positions of Glx, in addition to the methyl side-chain moieties. This brings the total of amino acid side-chain positions that can be simultaneously probed using a single ^{1}H dispersion experiment to 16. The utility of the approach is demonstrated with an application to the four-helix bundle colicin E7 immunity protein, Im7, which folds via a partially structured low populated intermediate that interconverts with the folded, ground state on the millisecond time-scale. The extracted ^{1}H chemical shift differences at side-chain positions provide valuable restraints in structural studies of invisible, excited states, complementing backbone chemical shifts that are available from existing relaxation dispersion experiments.



INTRODUCTION

The development of Carr–Purcell–Meiboom–Gill relaxation dispersion NMR spectroscopy[1,2] has opened up a new avenue for the study of low populated and transiently formed ('excited') protein states that interconvert with a highly populated ('ground state') conformation on the millisecond time-scale.[3,4] In addition to providing rates of exchange and the fractional populations of each of the participating states, absolute values of chemical shift differences between ground and excited conformers are also obtained that form the basis for calculating the shifts of the excited state(s) once signs are available from other experiments.[5,6] It is now possible to measure backbone ^{15}N,[7,8] ^{1}H^{N},[9] ^{13}C$^{\alpha}$,[10] ^{13}CO,[11,12] and ^{1}H$^{\alpha}$[13] chemical shifts of 'invisible' excited states as well as anisotropic restraints, such as residual dipolar couplings[14−16] or residual chemical shift anisotropies,[17] that orient internuclear bond-vectors or tensorial frames with respect to the external magnetic field. These restraints, obtained exclusively from CPMG-based relaxation dispersion experiments, can then be used in a variety of structure calculation programs[18−20] to produce atomic resolution models of 'invisible' protein states.[21−24]

With the development of CPMG dispersion experiments for measuring backbone chemical shifts reaching a 'mature' phase, attention has now turned to side-chain probes of structure. Experiments for quantifying chemical exchange at ^{13}C$^{\beta}$ carbons,[25] methyl ^{13}C[26] and ^{1}H[27,28] positions, side-chain ^{13}CO groups of Asx and Glx,[29−31] side-chain ^{15}N spins in Asn and Gln,[32] and the $^{15}NH_3^+$ moieties of Lys[33] have been developed that supplement the backbone data. The majority of the side-chain experiments to date have focused on heteronuclear probes of chemical exchange and with the exception of ^{1}H methyl relaxation dispersion experiments[27,28] there has been little development of ^{1}H methodology specifically tailored for side-chain positions. Extending ^{1}H CPMG dispersion to these sites would represent an important advance for a number of reasons. First, a significantly larger subset of residues would be available to serve as reporters of side-chain conformation, leading to improvements in the structures of proteins that are currently produced using chemical shift restraints. Second, ^{1}H chemical shifts are a particularly powerful reporter of structure[34]—including interactions between residues distal in primary sequence—because they are influenced by ring currents in aromatic groups. In the case of CPMG relaxation dispersion studies this type of data assumes an even more important role than in 'traditional' NMR structural investigations because it takes the place of NOE restraints that currently cannot be measured for excited state conformers.

Prior to the development of ^{1}H CPMG experiments for studies of conformational exchange at side-chain positions in proteins it is prudent to carefully evaluate labeling schemes. For example, artifacts must be avoided that would normally arise during the CPMG pulse train from evolution of ^{1}H magnetization due to ^{1}H–^{1}H scalar couplings in a fully protonated

Received: November 14, 2011
Published: February 2, 2012

   dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

spin network.[13] One approach is to produce proteins with 'designer amino acids' that have the exact labeling required, whereby proton spins are isolated by deuterating the positions adjacent to the site of interest. The SAIL method pioneered by Kainosho and co-workers[35] would be an optimal strategy in this regard, although likely expensive. An alternative approach, based on a labeling scheme that was originally introduced by Cowburn and co-workers for NMR structural studies of proteins,[36] produces fractionally deuterated molecules via expression in D_2O using $\{^{13}C,^{1}H\}$-glucose as the sole carbon source. Such a labeling procedure has been suggested recently by Mulder and co-workers for ^{1}H CPMG studies of methyl groups in proteins[27] since a large population of $^{13}CHD_2$ methyl probes attached to highly deuterated carbon sites (with the exception of Thr and Ile $C^{\delta 1}$) is produced. Guo and Tugarinov have also exploited this methyl labeling pattern in their studies of the 82 kDa enzyme malate synthase G.[37]

Here we show that the D_2O, $\{^{13}C,^{1}H\}$-glucose growth scheme generates proteins with many additional side-chain sites, beyond the methyl positions, that are suitable for analysis via ^{1}H CPMG relaxation dispersion. These include the β positions of Asx, Cys, Ser, His, Phe, Tyr, and Trp as well as the γ positions of Glx, corresponding to an additional 10 residues that can be studied and to approximately 40% of the amino acid content in a 'typical' bacterial protein.[38] In what follows, we first provide a brief analysis of the metabolic pathways responsible for amino acid biosynthesis that gives insight into how the labeling patterns are produced using D_2O, $\{^{13}C,^{1}H\}$-glucose protein expression. Second, an experimental scheme is presented for measuring ^{1}H CPMG relaxation dispersion profiles that involves recording a pair of data sets that are subsequently manipulated to produce spectra with correlations derived from Asx, Glx and aromatic amino acids or from all of the remaining residues. In this way all of the relevant ^{1}H side-chain sites, including those from methyl groups, can be probed simultaneously with no increase in measuring time. Finally, the utility of the methodology is demonstrated on the four helix bundle 87-residue colicin E7 immunity protein, Im7, that folds via a partially structured low populated intermediate on the millisecond time-scale.[29,39]

MATERIALS AND METHODS

Protein Sample Preparation. Isotopically enriched samples of the wild-type Im7 protein were expressed as described previously[29] (see Supporting Information for details) using 1 g/L of $^{15}NH_4Cl$ and 3 g/L of $\{^{13}C,^{1}H\}$-glucose as the sole nitrogen and carbon sources in 97% recycled D_2O. NMR samples (~ 1.4 mM in protein) were dissolved in 50 mM potassium phosphate buffer, pH 6.6, 0.02% azide, 100% D_2O or 95% H_2O/5% D_2O.

Details of Pulse Scheme of Figure 2. Narrow and wide solid vertical bars denote 90° and 180° rectangular pulses, respectively, of phase x unless indicated otherwise, while open bars correspond to composite 180 pulses[40] ($90_x180_y90_x$). All rectangular ^{1}H and ^{13}C pulses are applied with maximum power, with the exception of the ^{1}H refocusing pulses that comprise the CPMG pulse train and the flanking 90° pulses that are applied with a 25 kHz field. Initially, the ^{1}H carrier is centered on the water signal, while the ^{13}C transmitter is positioned at 38.5 ppm (aliphatic $^{13}C^{\beta/\gamma}$ region of Asx/Glx). At point a the proton carrier is moved to 2 ppm and returned to the position of the water line at point b just prior to acquisition. The deuterium channel is centered at 2 ppm with ~500 Hz WALTZ-16 decoupling[41] during the CT t_1 element; flanking 90° pulses are applied with a 1.6 kHz field. The symmetrically shaped ^{13}C pulses (indicated by 'ali') are of the REBURP variety,[42] are centered at 40.5 ppm with durations of 430 (315) μs at 11.7 (18.8 T) and cover the aliphatic region (10−71 ppm). IBURP-2 pulses[42] (indicated by 'aro/CO') are centered at 147.5 ppm (109 ppm phase modulation of the carrier centered at 38.5 ppm) and applied for durations of 430 (11.7 T) or 260 (18.8 T) μs, with the second of the two time-reversed, as illustrated. These invert over a range extending from 110 to 185 ppm, covering both side-chain ^{13}CO (Asx/Glx) and $^{13}C^{\gamma}$ (aromatic) carbons. Decoupling during acquisition is achieved with a 2.0 (3.0) kHz GARP-1 field[43] at 11.7 (18.8 T). Suppression of $^{13}CH_2$ and $^{13}CH_2D$ moieties is achieved using either the $90_x90_{\phi 4}$ element or a single 90° purge pulse (the gray 90° ^{1}H pulse is removed), the latter attenuating residual signal from $^{13}CH_3$ isotopomers (see Supporting Information). Delay times are $\tau_a = 1.85$ ms $\approx 0.25/J_{CH}$, $\tau_b = 3.8$ ms $= 0.5/J_{CH}$ (for methylene groups), TC = 13.6−14.0 ms $\approx 0.5/J_{CC}$ (for aliphatic carbons), $\delta = 0.4$ ms, $\tau_{eq} = 5$ ms $\geq (2-3)/k_{ex}$. The value T $(= 4n\cdot\tau_{cpmg})$ is chosen such that approximately 1/2 of the signal has decayed relative to $n = 0$ (40 ms for the applications considered here), where n can be any integer. The phase cycle is $\phi_1 = x,-x$, $\phi_2 = 2(x),2(-x)$, $\phi_3 = x$, $\phi_4 = x,-x$, $\phi_5 = x,-x,y,-y$, $\phi_6 = x$ and $\phi_{rec} = x,2(-x),x$. Quadrature detection in F_1 is achieved with the enhanced sensitivity gradient selection approach,[44] whereby a pair of data sets are recorded for each t_1 value corresponding to (ϕ_6, gradient 7) and ($-\phi_6$, gradient 7). Axial peaks are shifted to the edge of the spectrum by inverting ϕ_3 and ϕ_{rec} with each t_1 increment.[45] Gradient strengths in G/cm (lengths in ms) are: 0 = 16 (0.5), 1 = 10 (0.25), 2 = 25 (1.0), 3 = 20 (0.5), 4 = 30 (0.15), 5 = 16 (0.5), 6 = −20 (0.8), 7 = 7.544 (0.15).

Recording and Fitting of Experimental Data. Side-chain proton relaxation dispersion data sets were recorded at static magnetic fields of 11.7 T and 18.8 T using the pulse scheme of Figure 2. Data sets recorded at 25 °C, 11.7 (18.8) T were obtained with 21 (23) ν_{CPMG} values ($T = 40$ ms) ranging from 25 to 2000 Hz, including three ν_{CPMG} repetitions for error analysis. Each 2D spectrum (11.7 T) was recorded with 8 transients, (46,512) complex points in (t_1,t_2), corresponding to acquisition times of (26.9 ms, 64 ms), and a delay between scans of 2.5 s (net measurement time of 65 min/spectrum). Data sets were obtained at 18.8 T with 4 transients and comprised (84,768) complex points, acquisition times of (28 ms, 64 ms). A relaxation delay of 2.5 s was used, giving rise to acquisition times of 65 min/spectrum. Similar acquisition parameters were used for data recorded at 10 °C.

NMR data were processed and analyzed using the NMRPipe suite of programs[46] and Sparky[47] with peak volumes extracted using the program FuDa (http://pound.med.utoronto.ca/software.html). CPMG dispersion profiles were fit to a two-site exchange model, $A \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} B$, as described previously[29] using in-house written software, CATIA (http://pound.med.utoronto.ca/software.html). The use of the simplest of exchange models has been shown previously to be sufficient to describe the exchange process, corresponding to the interconversion between a metastable folding intermediate and the native conformation of Im7.[29] The analysis of the dispersion profiles was performed as described in detail in a previous publication.[29] Briefly, dispersion profiles were first fitted on a per-residue basis and residues selected for determining the global exchange parameters, p_B, the population of the excited state, and $k_{ex} = k_{AB} + k_{BA}$, using criteria described in reference 29. In total, dispersion profiles from 5 H^{β} protons from aromatic residues, 17 H^{β}/H^{γ} from Asx/Glx, and 26 methyls were retained at this stage (pH 6.6, D_2O, 25 °C), while dispersions from 12 H^{β}/H^{γ} from Asx/Glx and 29 methyl groups were fitted to extract parameters for the data recorded on the sample at pH 6.6, H_2O, 10 °C. Subsequently, the (k_{ex}, p_B) parameters were fixed to those from the global analysis of the selected residues and every dispersion profile fitted to extract chemical shift differences between ground and excited states ($\Delta\omega$, rad/s or $\Delta\varpi$, ppm), with errors determined from a bootstrap procedure.[48]

Measurement of Rotating Frame $^{1}H-^{1}H$ Cross-Relaxation Rates, σ. Values of σ were obtained as described by Ishima et al.,[49] by recording ^{1}H $R_{1\rho}$ rates for $^{13}CHD_2$ and ^{13}CHD groups by including a ^{1}H spin-lock element (10 kHz) of variable duration that is placed either (i) at the start of the scheme of Figure 2, $R_{1\rho}^{A}$ or (ii) immediately at the end and therefore after the t_1 evolution period, $R_{1\rho}^{B}$,

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

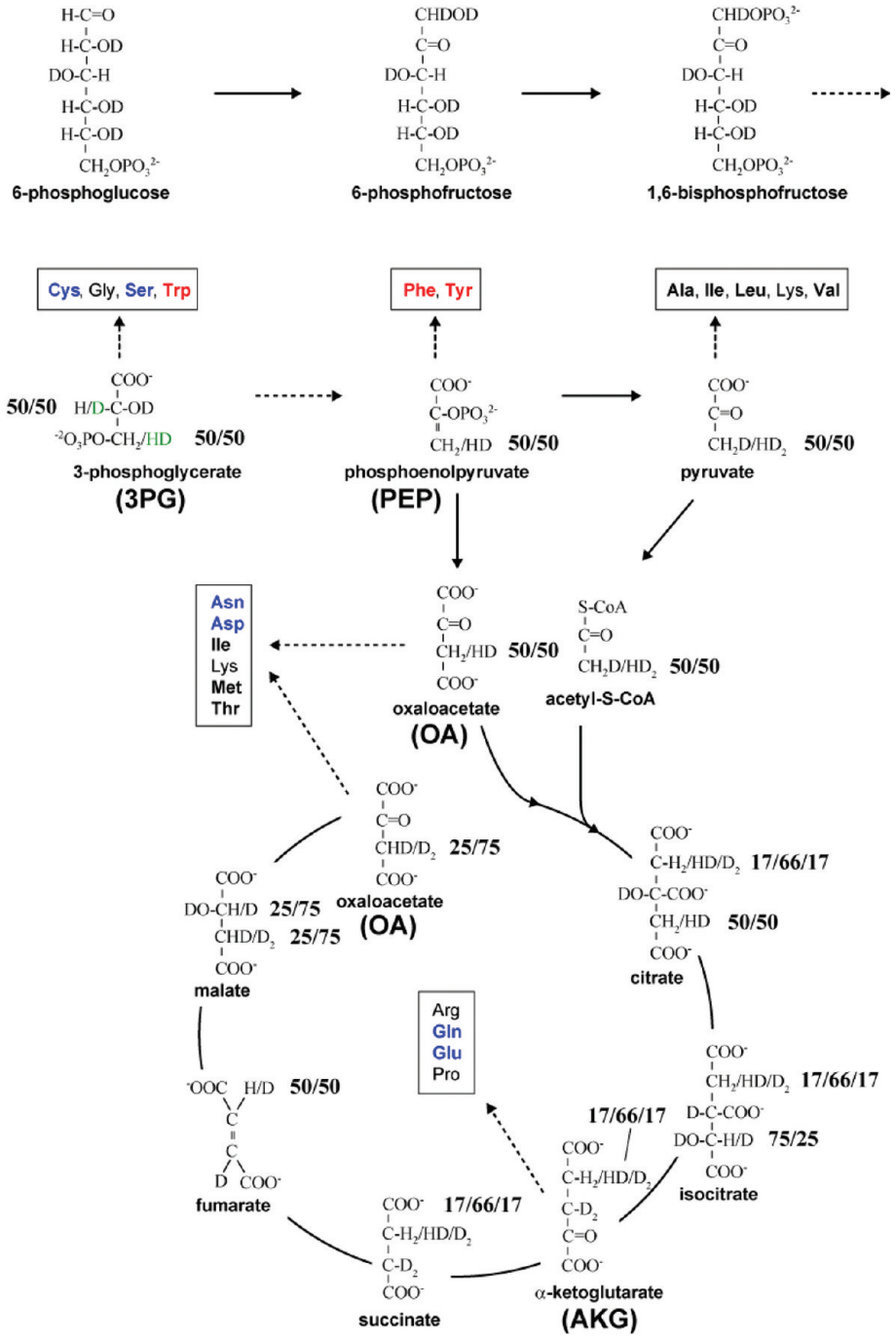


Figure 1. Biosynthesis pathways highlighting key steps in the production of many of the amino acid side chains as well as the expected H/D isotopomer distributions of metabolic precursors generated from {^{13}C,^{1}H}-glucose, D_2O. Residues highlighted in blue (Asx, Cys, Glx, Ser), red (aromatic), or bold (methyl-containing) are those that contain side-chain ^{1}H probes that are available for study using the CPMG relaxation dispersion pulse scheme of Figure 2 (also His). Dashed (solid) arrows connect compounds that are related via multiple (single) steps. The isotopomer distributions at distinct carbons of a given compound are not correlated with the exception of those for 3-phosphoglycerate where the distribution is (H,H_2) and (D,HD), the latter indicated in green. Abbreviations used for a number of the precursors are indicated in parentheses.

using a version of the experiment of Figure 2 that is not sensitivity enhanced. Ishima et al. have shown that $R_{1\rho}^{A} > R_{1\rho}^{B}$ and that the difference gives the value of σ.[49] Alternatively, rotating frame cross-relaxation rates can be estimated by recording the decay of ^{1}H,^{13}C longitudinal order, $R(I_ZC_Z)$ and ^{13}C longitudinal magnetization, $R_1(^{13}C)$, with $\sigma \approx 2\{R(I_ZC_Z) - R_1(^{13}C)\}$ (the factor of two accounts for the fact that the magnitude of rotating frame cross-relaxation rates are 2-fold larger than longitudinal rates in the limit of slow molecular tumbling[50]). Similar values were obtained from both sets of measurements.

RESULTS AND DISCUSSION

The goal of this work is to develop a simple approach for quantifying conformational exchange at side-chain positions in proteins using ^{1}H CPMG relaxation dispersion NMR. This involves, first, choosing a suitable labeling scheme that facilitates measurement of dispersion profiles that are free from artifacts and second, exploiting the labeling strategy in the design of the appropriate pulse scheme for carrying out the measurements. The success of any ^{1}H-based CPMG experiment is limited by the

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

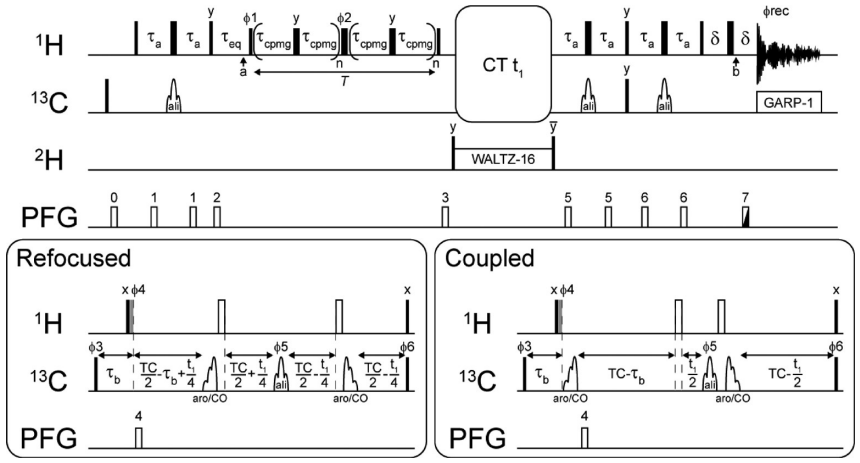


Figure 2. Pulse scheme for measuring ^{1}H side-chain CPMG relaxation dispersion profiles for uniformly ^{13}C-labeled, partially deuterated proteins. Details are given in Materials and Methods.

extent to which the ^{1}H spin(s) of interest can be isolated from adjacent protons that potentially form scalar coupling partners. In the case of a fully protonated protein, for example, the application of successive CPMG refocusing pulses leads to magnetization transfer between scalar coupled homonuclear (^{1}H) spins with an efficiency that depends on the CPMG pulse rate, $\nu_{CPMG} = 1/(4\tau_{CPMG})$, where $2\tau_{cpmg}$ is the time between successive refocusing pulses in the CPMG pulse train.[13] Thus, dispersion profiles are 'contaminated' by the effects of homonuclear *J*-evolution leading to modulations that are not easily separated from those due to chemical exchange.

Previously, we had shown that it is possible to record high-quality $^{1}H^{\alpha}$ CPMG relaxation dispersion profiles of proteins produced with high levels of deuteration at side-chain positions and with approximately 50% protonation at C^{α} sites.[13] Such proteins were obtained by expression in 50% H_2O/50% D_2O media using {^{13}C,^{2}H}-glucose as the sole carbon source. Here we wish to exploit an analogous labeling strategy for extending the measurements to a number of side-chain ^{1}H positions in proteins. We have chosen to focus on a very simple approach whereby proteins are produced using 100% D_2O as solvent and {^{13}C,^{1}H}-glucose as the carbon source, a labeling scheme that was introduced previously in structural studies of moderately sized molecules[36] and used subsequently to produce proteins with $^{13}CHD_2$-labeled methyl groups for ^{1}H CPMG dispersion studies.[27] We show below, first by consideration of the biosynthetic pathways for the amino acids and subsequently by experiment, that a significant number of isolated side-chain ^{1}H probes become available in proteins generated by this approach, in addition to those associated with methyl groups that have been exploited previously.[27] These include $^{1}H^{\beta}$ of Asx, Cys, and Ser, $^{1}H^{\gamma}$ of Glx, and $^{1}H^{\beta}$ of aromatic amino acids—the 10 residues that are the primary focus of what follows. Combined with the methyl probes of Ala, Ile, Leu, Met, Thr, and Val this brings the number of distinct side-chain types to 16 that can be used to study millisecond time-scale exchange via ^{1}H CPMG transverse relaxation methods.

A Labeling Scheme for the Measurement of Side-Chain ^{1}H Dispersion Profiles. Despite a bewildering number of side-chain positions in the 20 amino acids, the isotopomer composition of many residues in proteins generated from the D_2O/{^{13}C,^{1}H}-glucose approach[36] can, surprisingly, be predicted from consideration of only three biosynthetic pathways, including glycolysis, the TCA cycle and the pentose phosphate pathway.[51] In what follows we describe the biosynthesis of Asx, Cys, Ser, Glx and the aromatic amino acids (aliphatic positions) that are considered in some detail here. For completeness we also include a brief description of the biosynthesis of the methyl-containing amino acids. Throughout the discussion the reader is referred to Figure 1 where the main steps in glycolysis and the TCA cycle are highlighted, along with the levels of protonation and deuteration at key positions. It is worth noting that the ^{1}H/^{2}H isotopomer composition at each carbon is independent of other sites, except for 3-phosphoglycerate (3PG), discussed below. The first important step in glycolysis is the isomerization of 6-phosphoglucose to 6-phosphofructose, resulting in the addition of a deuteron at position 1 and the elimination of a proton at position 2 (for ease of description all numbering of molecules is from top to bottom that does not necessarily coincide with chemical convention). The two molecules of 3PG that subsequently result from the cleavage of 1-6-bisphosphofructose will thus have different isotopomer compositions at positions 2 and 3, either H,H_2 or D,HD, respectively. In turn, phosphoenolpyruvate (PEP) that follows 3PG will have H_2 or HD labeling at position 3. Since C3 of either 3PG or PEP is the precursor for the β-position of the amino acids Cys, Ser, Trp (3PG) and Phe, Tyr (PEP) these amino acids are predicted to have 50% H_2 and 50% HD at C^{β}. The addition of a deuteron from solvent (D_2O) results in pyruvate, and its isotopomer composition at position 3 is thus predicted to be 50% H_2D and 50% HD_2. The methyl group of pyruvate is transferred to the corresponding methyl positions of Ala, $Ile^{\gamma 2}$, Leu, and Val so that the methyls of these residues are labeled as CH_2D (50%) or as CHD_2 (50%). In addition to serving as the precursor for Phe and Tyr, PEP is carboxylated to form oxaloacetate (OA), and pyruvate is decarboxylated en route to becoming acetyl-S-CoA. These two molecules combine to form citrate, the starting point of the TCA cycle. In the process one of the hydrogens (either H or D) of acetyl-S-CoA is eliminated so that the corresponding position 2 in citrate will have the predicted isotopomer composition of 17% H_2, 66% HD, and 17% D_2 (see Figure 1). Through a series of reactions citrate is converted to α-ketoglutarate (AKG). It is noteworthy that both hydrogen atoms at position 3 of this molecule are derived from solvent and will thus be 100% D_2, while those at position 2 derive from

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

C2 of citrate, thereby retaining the 17% H_2, 66% HD, and 17% D_2 distribution.

AKG is the precursor of Arg, Gln, Glu, and Pro, and hence these amino acids will be 100% D_2 at the β-position and are expected to be labeled as 17% H_2, 66% HD, and 17% D_2 at the γ-position. Subsequently, AKG is decarboxylated to succinate, which, in turn, is dehydrogenated to yield fumarate. In this process one hydrogen each at positions 2 and 3 of succinate is eliminated so that fumarate will have an isotopomer composition of 50% H/50% D and 100% D at positions 2 and 3, respectively. However, because fumarate is a symmetric compound, positions 2 and 3 are equivalent so that, when D_2O is added, the resulting molecule malate and its product following oxidation, OA, will have an isotopomer composition 25% HD and 75% D_2 at position 3. It is noteworthy that both hydrogen atoms at this position are derived from either acetyl-S-CoA or solvent and are lost before AKG is formed in subsequent rounds of the TCA cycle. It is thus sufficient to consider only the first pass.

The β-position of the amino acids Asn, Asp, Met, and Thr as well as the γ1-position of Ile derive from position 3 of OA and since OA may be formed either by carboxylation of PEP or via the TCA cycle it is necessary to know the relative fluxes of these two pathways to predict the isotopomer composition at these sites. In a previous study we found that 70% of OA was formed via the TCA cycle[13] and assuming the same value here, the predicted isotopomer composition at the β-position for Asn, Asp, and Met is 15% H_2, 33% HD, and 52% D_2. Since one of these hydrogen atoms is eliminated in the biosynthesis of Thr, this amino acid will have a 32% H/68% D composition at the β-position.

It is useful to summarize at this point the expected labeling patterns of key positions on the basis of the analysis presented above. With regards labeling at β sites, Asx is predicted to be 15% H_2/33% HD/52% D_2, Glx 100% D_2, and Ser/Cys 50% H_2/50% HD while the γ positions of Glx are expected to be 17% H_2/66% HD/17% D_2. Further, the β positions of the aromatic residues Phe, Trp, and Tyr are 50% H_2/50% HD; the isotopomer composition of His is more difficult to predict because there are a number of different contributing pathways. Because protein synthesis is carried out using D_2O solvent, the α positions are expected to be 100% deuterated. Indeed, on the basis of a comparative analysis of ^1H,^{13}C correlation spectra of protein samples prepared using {^{13}C,^1H}-glucose and ~97% D_2O solvent (see Materials and Methods) we estimate that the level of protonation at the α position in protein prepared for ^1H relaxation dispersion studies to be less than 4% on average (Table 1, see Supporting Information for details), in agreement with what Mulder and co-workers have found previously.[52] In addition, further experiments establish that the isotopomer distribution at the β-methylene position of Glx, expected to be CD_2, is on average approximately 14% CHD, 86% CD_2 (Supporting Information, Table 1). The β-sites of Asx, Cys, Ser and the aromatic amino acids, along with γ-sites in Glx are thus well isolated from protons (i.e., low probability of protonation on adjacent carbons). Moreover, these methylene sites are predicted to contain sizable fractions of the HD isotopomer that can be selected in NMR experiments (see below), thereby avoiding complicating effects in ^1H CPMG experiments from magnetization transfer between geminal ^1H pairs. The predicted isotopomer distributions for the residues discussed above are summarized in Table 1A and the experimentally determined values are listed in Table 1B (see Supporting Information for details). Overall, the agreement between predicted and experimentally determined CHD isotopomer populations is very reasonable for Asx, Glx (at both β and γ positions), Phe and Tyr, while there are differences for Ser and Trp. These discrepancies may reflect a kinetic isotope effect whereby the protonated form of 3PG (see Figure 1) is preferred over the fractionally deuterated precursor (green hydrogens in 3PG) by phosphoglycerate dehydrogenase, an enzyme in the biosynthetic pathway of Ser and Trp that oxidizes the 2 position of 3PG to a ketone. Dehydrogenase enzymes such as the one mentioned above are known to have significant kinetic isotope effects.[53,54] A second contributing factor to the higher protonation levels is that 3PG is produced by the pentose phosphate pathway as well, via the glycolytic intermediate glyceraldehyde-3-phosphate (GAP) that is fully protonated at the position corresponding to Cβ of Ser and Trp.[51] This leads to a dilution of the CHD isotopomer. It is tempting to think of supplementing the growth medium with an inhibitor of the enzyme transketolase that catalyzes the production of GAP in the pentose phosphate pathway. However, transketolase also plays a role in the biosynthesis of precursors for aromatic side-chains so that addition of such an inhibitor could very well lead to slow growth rates and ultimately low protein yields. We have not, therefore, attempted to 'manipulate' the growth

Table 1. [b]Predicted and Experimental Isotopomer Composition at Selected Non-methyl-Containing Side-Chain Positions in Amino Acids Synthesized by *E. coli* in M9 Medium Supplemented with Protonated Glucose and 100% D_2O Solvent (predicted)[a]

A. predicted		
amino acid	β	γ
Asx	15% H_2, 33% HD, 52% D_2	
Cys	50% H_2, 50% HD	
Glx	100% D_2	17% H_2, 66% HD, 17% D_2
His[c]		
Phe	50% H_2, 50% HD	
Ser	50% H_2, 50% HD	
Trp	50% H_2, 50% HD	
Tyr	50% H_2, 50% HD	
B. experimental[d]		
amino acid	β	γ
Asx (N = 9)	12 ± 1% H_2, 26 ± 7% HD	
Cys (N = 0)		
Glx(N^β=14), (N^γ = 17)	0% H_2, 14 ± 6% HD	8 ± 3% H_2, 47 ± 15% HD
His (N = 2)[e]	0% H_2, 21 ± 17% HD	
Phe/Tyr (N = 11)	10 ± 2% H_2, 48 ± 6% HD	
Ser (N = 7)	64 ± 7% H_2, 28 ± 10% HD	
Trp (N = 2)	67 ± 1% H_2, 25 ± 5% HD	

[a]Details for methyl containing side chains have been published previously.[52] [b]A tabulated value of x% HD means that x/2% of the methylene groups are CHD with protonation at one of the two prochiral positions, while x/2% of the CHD methylenes are protonated at the other position. [c]Difficult to predict because several pathways contribute to this position. [d]Measurements utilized an Im7 protein sample prepared with 97% D_2O, so that it is expected that slightly higher deuteration levels would be possible using >99% D_2O. N refers to the number of peaks in $^{13}C-^1H$ correlation spectra that have been quantified to arrive at the fractional populations listed. Experimental details are given in Supporting Information. [e]Peaks are weak, likely due to residual broadening at this position, and quantification is unreliable.

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178–3189

conditions to obtain higher yields of the CHD isotopomer at Ser and Trp positions. A detailed description of the conditions used for Im7 protein overexpression is given in Supporting Information.

In addition to the positions indicated in Table 1, many of the methyl sites in proteins produced using {^{13}C,^{1}H}-glucose/D_2O are also available for ^{1}H CPMG dispersion analysis, as originally demonstrated by Mulder and co-workers.[27] This includes the methyls of Ala, Ile$^{\delta 2}$, Leu and Val where the fractional populations of the CHD_2 isotopomer are approximately 50%.[52] Notably, the positions adjacent to these methyls are highly deuterated so that their ^{1}H spins are well isolated, as is the case for Met since the $^{13}C^{\epsilon}$ methyl group is connected to the side-chain C^{γ} through a sulfur. By contrast, measurement of methyl ^{1}H transverse relaxation rates for Thr and Ile$^{\delta 1}$ by a CPMG pulse-train is complicated by the fact that the adjacent carbons are not completely deuterated (predicted 32% H/68% D for Thr β and 32% HD/68% D_2 for Ile $C^{\gamma 1}$, see above and reference 52), leading to modulation of magnetization from three-bond ^{1}H−^{1}H scalar couplings. In these cases, however, we have found that it is possible to resolve separate methyl correlations that arise from ^{2}H isotope shifts due to adjacent isotopomers and to 'focus' only on the peak that reflects the fully deuterated adjacent carbon position, thus avoiding this problem (see below). Mulder et al. have designed pulse sequences that specifically select for methyl groups that are attached to fully deuterated carbons, circumventing the problem at the expense of either lower resolution or decreased sensitivity in spectra.[27]

Measurement of Side-Chain ^{1}H CPMG Relaxation Dispersion Profiles. Figure 2 illustrates the pulse scheme that has been developed to record side-chain ^{1}H dispersion curves for proteins expressed in D_2O, {^{13}C,^{1}H}-glucose media. The magnetization 'flow' during the course of the sequence can be summarized as follows

$$^{1}\mathrm{H} \rightarrow \mathrm{CPMG}(T) \rightarrow {}^{13}\mathrm{C}(\mathrm{CT}t_1) \rightarrow {}^{1}\mathrm{H}(t_2)$$

where CPMG(T) is the CPMG element during which a variable number of ^{1}H 180° pulses are applied. As with our previous ^{1}H CPMG scheme for measuring dispersion profiles in $^{13}CHD_2$-labeled methyl groups[28] we have chosen not to add an additional (P-) element at the midpoint of the CPMG period that averages out the imbalance in relaxation rates of ^{1}H in-phase (I_{TR}) and antiphase (with respect to the ^{13}C, $2I_{TR}C_Z$) magnetization.[7] Instead, the imbalance is taken into account explicitly using a two-spin scalar coupled basis (^{1}H, ^{13}C) in calculating the dispersion profiles during the fitting procedure with the program CATIA, as described previously.[14] It should be noted that the effects of such differential relaxation are, in any event, small. As described previously, $R_2(I_{TR}) - R_2(2I_{TR}C_Z) \approx R_1(^{13}\mathrm{C})$, where $R_1(^{13}\mathrm{C})$ is the ^{13}C longitudinal relaxation rate.[28] Values of $R_1(^{13}\mathrm{C})$ have been measured for the Im7 protein, 25 °C, 11.7T, with $R_1(^{13}\mathrm{C}) \approx 1\ \mathrm{s}^{-1}$ for $^{13}CHD_2$ methyl groups and between 2 and 3 s^{-1} for ^{13}CHD methylenes; values $<1\ \mathrm{s}^{-1}$ are predicted for methylene groups at 18.8 T based on measurements at 11.7 T. Simulations on a nonexchanging spin system where $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}})$ is expected to be flat establish that this differential relaxation leads to small deviations from flat dispersion profiles, with $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}} = 0) - R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}} \rightarrow \infty) < 1.5\ \mathrm{s}^{-1}$ (11.7 T) and $0.5\ \mathrm{s}^{-1}$ (18.8 T). The benefit of omitting the P-element is that the lowest ν_{CPMG} frequency that can be used is a factor of 2 smaller than otherwise possible, so that the initial dispersion regime can be more thoroughly mapped, an advantage in applications involving systems with moderately slow exchange rates[55] (see below) and for residues with large chemical shift differences between exchanging states $|\Delta\varpi|$, irrespective of the exchange time scale.

The CPMG relaxation period is followed by a constant time interval, CT t_1, during which ^{13}C chemical shift evolution proceeds. Its duration is set to $\sim 1/J_{CC}$, where J_{CC} is the one-bond aliphatic ^{13}C−^{13}C coupling constant (~35 Hz), since the sample is fully ^{13}C labeled. During this interval the appropriately labeled methylene(methyl) group, ^{13}CHD($^{13}CHD_2$), is selected. The choice of only singly protonated methylene groups eliminates artifacts in dispersion profiles that would arise from magnetization evolution due to the large geminal ^{1}H−^{1}H coupling, and it is further noteworthy that selection of groups with only a single proton significantly attenuates ^{13}C transverse relaxation rates, improving the sensitivity of the experiment. Selection against the undesired isotopomers is achieved by allowing ^{13}C magnetization to evolve due to the one-bond ^{1}H−^{13}C scalar coupling for a duration of $\tau_b = 1/(2J_{HC})$ that is followed by a ^{1}H 180° pulse in alternating scans. The net effect is that magnetization from $CH_2(CH_2D)$ groups, but not CHD(CHD_2), is inverted in successive scans, so that addition of signal leads to selection of the desired labeling pattern. Key to the removal of undesired isotopomers in the filter used here is the uniformity of one-bond J_{HC} values for the residues of interest (except Ser and Met). Values measured for the Im7 protein are 130.8 ± 1.9 Hz (Asp, β), 127.0 ± 0.9 (Glx, γ), 130.7 ± 2.7 (aromatic residues, β), 146.9 ± 0.6 (Ser, β); one-bond J_{HC} values have been reported previously for $^{13}CHD_2$ methyl groups,[56] ranging from 124.4 Hz for Leu to 129.4 Hz for Ala, with Met the only substantial outlier where a coupling of 138 Hz is measured. It is clear that a compromise value of τ_b must be chosen, and we have used a value of 3.8 ms, corresponding to $J_{HC} \approx 130$ Hz in what follows here. The $90_x90_{\phi 4}$ filter described above does not suppress methyl groups of the $^{13}CH_3$ isotopomer variety and relative intensities of the $^{13}CH_3$:$^{13}CHD_2$ correlations can be as much as 0.2, corresponding to approximately 7% $^{13}CH_3$ isotopomer relative to $^{13}CHD_2$. A factor of approximately 2-fold improvement in $^{13}CH_3$ suppression is achieved by replacing the $90_x90_{\phi 4}$ pulse pair with a single 90° purge pulse that works equally well for suppression of $^{13}CH_2$ and $^{13}CH_2D$ moieties (see Supporting Information). Better suppression could be achieved at the expense of a longer pulse scheme but we have opted for the simplest (and most sensitive) approach here.

In addition to selection for the appropriate label at the methylene/methyl positions the CT element also facilitates selection of methylenes from Asx, Glx and the aromatic amino acids on the basis of the distinct carbon chemical shifts of their C^{γ} (Asx, aromatics) and C^{δ} (Glx) carbons.[57] This is achieved by recording a pair of spectra where the $^{13}C^{\beta}$−$^{13}C^{\gamma}$ coupling (Asx, aromatics) or $^{13}C^{\gamma}$−$^{13}C^{\delta}$ coupling (Glx) is either refocused ('Refocused' scheme of Figure 2) or allowed to evolve ('Coupled') through the use of band selective pulses that invert selectively the aromatic/carbonyl carbons or the aliphatic region of the spectrum. In the 'Coupled' spectrum $^{13}C^{\beta}$−$^{13}C^{\gamma}$/$^{13}C^{\gamma}$−$^{13}C^{\delta}$ scalar evolution proceeds for a duration of $2(T_C - \tau_b)$ that is tuned to approximately $1/(J^{\mathrm{Asx,Aro}}_{C\beta-C\gamma})$, $1/(J^{\mathrm{Glx}}_{C\gamma-C\delta})$ (19.6 ms). Subtraction of the pairs of data sets recorded in this manner produces a spectrum with correlations from methylenes of aromatics and Asx/Glx residues. Interestingly, very weak ($^{13}C^{\alpha}$,$^{1}H^{\alpha}$) cross-peaks are also observed. These arise due to the small amount of protonation at the α-position (<5%) and the fact that the backbone fragment $^{1}H^{\alpha}$−$^{13}C^{\alpha}$−^{13}CO is selected in the

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

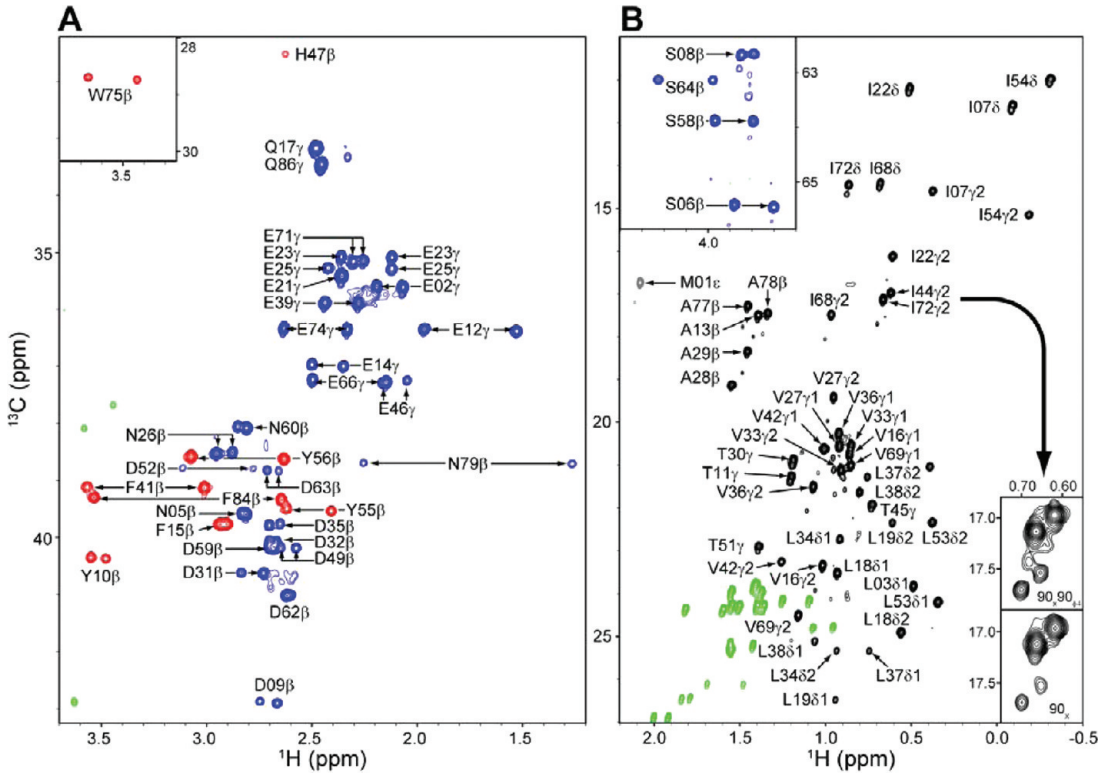


Figure 3. Spectra recorded using the pulse scheme of Figure 2, $T = 0$, by subtracting 'Refocused' and 'Coupled' data sets to generate a correlation map of selected methylene groups labeled as ^{13}CHD, A, (Asx β, Glx γ, aromatic β) or adding, B, to obtain, among other cross-peaks, those derived from Ser β (^{13}CHD) and methyl groups (^{13}CHD$_2$). M01ε (in B) is colored gray to indicate that its intensity is inverted relative to other methyl groups since its methyl carbon is not coupled directly to ^{13}C. Residues are color coded as blue (Asx,Glx,Ser), red (aromatics), and black (methyl groups), as in other figures. "Green" peaks derive from other side-chain positions that are not of interest in the present study. The lower insets in B show expanded regions (same contour levels) illustrating I44 and I72 methyl correlations from spectra that were obtained using either the $90_x90_{\phi 4}$ (above) or the 90_x (below) purge. The main figures are of data sets generated using a pulse scheme with the 90_x purge. Data were recorded on a sample of Im7 (25 °C, pH 6.6, D_2O) at 18.8 T.

'Coupled' − 'Refocused' difference spectrum in the same manner as Asx, aromatics, and Glx side chains since the selective aromatic/CO pulses of Figure 2 invert both backbone and side-chain carbonyl magnetization. Addition of the data sets produces a spectrum containing correlations from the remaining residues, including Cys, Ser, and the methyl-containing amino acids. In total millisecond time scale dynamics from 16 side-chains can be probed using this experimental scheme.

Finally, as described above, we have chosen to place the CPMG element prior to the t_1 evolution period for the application to the small Im7 protein considered here. It is thus possible to make use of the enhanced sensitivity approach whereby both cosine and sine modulated t_1 magnetization components are detected in each scan, increasing the sensitivity of the experiment by as much as $\sqrt{2}$, Figure 2. There is a potential draw-back to this approach, however, as described previously by Ishima and Torchia in connection with similar dispersion experiments for ^{1}H^{N} protons.[9] At issue is the fact that the efficiency of rotating frame cross-relaxation between the proton probe of interest and adjacent protons depends on the chemical shift differences between the cross-relaxing spins, increasing with higher ν_{CPMG} frequencies that scale down the shift differences. This leads, therefore, to a modulation of transverse relaxation rates (and hence affects dispersion curves) in a manner that is not related to the exchange process *per se*. We have measured the rotating frame cross-relaxation rates, σ, for ^{13}CHD$_2$ and ^{13}CHD moieties in Im7, 25 °C, as described in Materials and Methods, with values ~1 s^{-1} for methyls and methylenes of all amino acids with the exception of aromatic residues where $\sigma_{avg} \approx 2$ s^{-1} ($\sigma_{max} \approx 3$ s^{-1}). Simulations of dispersion profiles in the absence of exchange but including cross relaxation as well as differences between $R_2(I_{TR})$ and $R_2(2I_{TR}C_Z)$ rates indicate that most often errors in $R_{2,eff}(\nu_{CPMG})$ are on the order of 1 s^{-1} or less, extending to 2 s^{-1} in a number of cases, similar to intrinsic errors in relaxation rates obtained from repeat measurements. For the Im7 protein where dispersions are large (see below) such errors are tolerable. It is worth emphasizing, however, that the effects of cross-relaxation can be eliminated, at the expense of sensitivity, by placing the CPMG element after the t_1 evolution period, as described by Ishima and Torchia.[9]

An Application to the Im7 Protein. The colicin E7 immunity protein, Im7, is a small (87 residue) four-helix bundle protein that folds via a partially structured intermediate[58,59] (I). We have shown previously[29] that under the experimental conditions used here (pH 6.6, D_2O, 25 °C or pH 6.6, H_2O, 10 °C) the unfolded state is populated to such a small extent that dispersion profiles effectively report on the two-state exchange process linking the native structure (N) with I, $I \underset{k_{NI}}{\overset{k_{IN}}{\rightleftharpoons}} N$. Figure 3 shows spectra that were recorded using the pulse scheme of Figure 2, $T = 0$, subtracting the 'Refocused' and 'Coupled' data sets to generate a correlation map of selected methylene groups, A, (Asx β, Glx γ, aromatic β). Alternatively, if the data sets are added, the resultant spectrum comprises correlations from the remaining aliphatic moieties, including Cys β, Ser β,

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

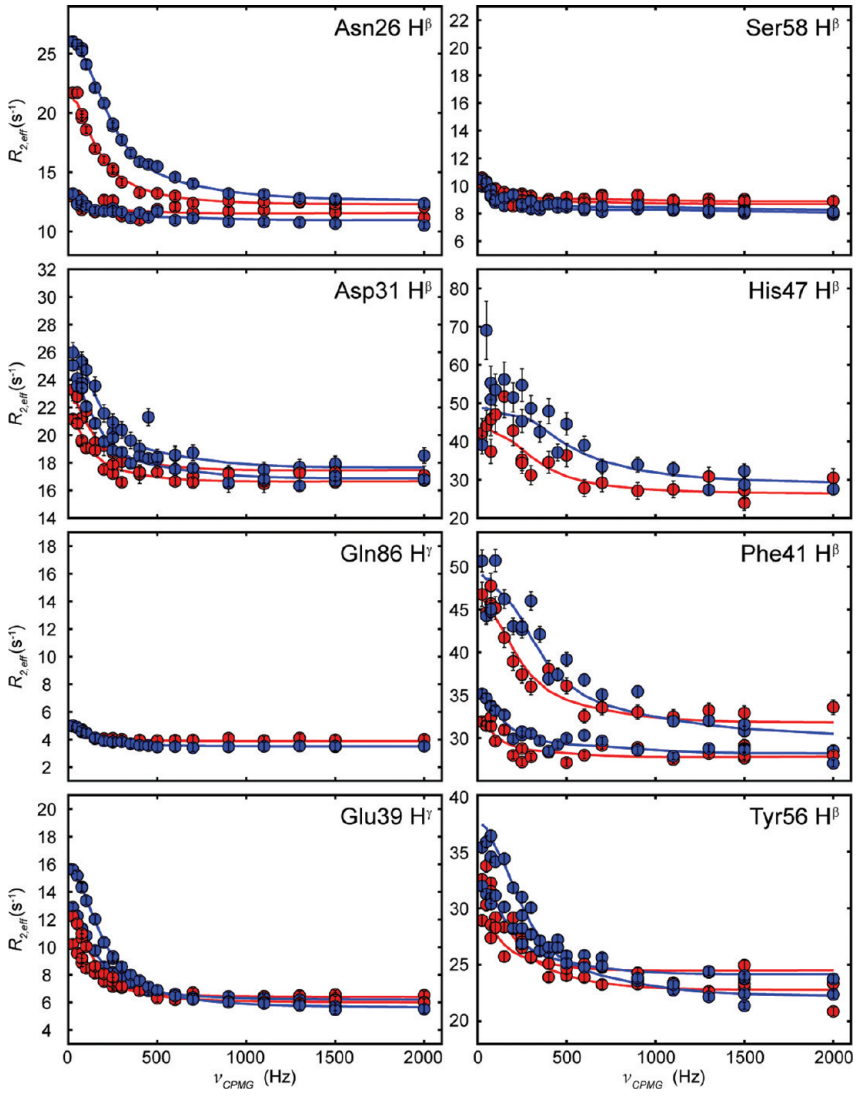


Figure 4. Representative side-chain ^{1}H CPMG relaxation dispersion profiles (Im7, 25 °C, pH 6.6, D_2O) highlighting data from eight of the ten nonmethyl-containing side chains that can be probed using the experiment of Figure 2. Shown are experimental data (circles) recorded at 11.7 T (red) and 18.8 T (blue) along with solid lines that correspond to the best fit of the profiles taken all together, as described in Materials and Methods. Both methylene proton dispersions are shown for a given residue, in cases where measurement of each was possible. Stereospecific assignments for the methylene protons are not available.

and the methyl groups that are valuable probes of exchange in proteins, Figure 3B.

Figure 4 shows representative CPMG relaxation dispersion curves for H^{β} and H^{γ} protons from a diverse set of side chains in Im7 recorded at 18.8 (blue) and 11.7T (red), circles, 25 °C, pH 6.6, D_2O. In addition, the best fit to the data using a two-site model of chemical exchange as described previously[29] is indicated by the solid lines, with the population of the intermediate state $p_I = 2.3 \pm 0.02\%$ and $k_{ex} = k_{IN} + k_{NI} = 910 \pm 10\ s^{-1}$ obtained from the global fit of all data including methyl ^{1}H dispersion profiles (48 pairs of dispersion profiles recorded at 11.7 T and 18.8 T, see Materials and Methods for details). Profiles for both methylene protons are shown in all cases (stereospecific assignments are not available currently), with the exception of His 47 where only a single weak (^{13}C,^{1}H) correlation is observed in spectra, Figure 3, and Gln 86 where H^{γ} chemical shifts are degenerate. While it is the case that very good fits of the dispersion data are obtained in general using a two-state exchange model (see also Figure 6) there are a number of exceptions, such as Phe 41 and especially His 47. Peaks derived from these residues are weak but perhaps more important is that Im7 exchange is highly pH dependent at neutral pH, pointing to a significant role for the two His side chains in the protein (His 40 and 47; correlations for His 40 are broadened to the point that they not observed). It may well be that His 47, located in the loop connecting helices 2 and 3 (see below), undergoes an exchange process that is more complex than two-state. Interestingly, the side-chain of Phe 41 (helix 2) projects toward the loop and may be 'affected' by the same complex dynamic process that modulates the chemical shifts of His 47. Figure 4 establishes that it is not always the case that similarly sized ^{1}H dispersions are measured for both methylene proton positions of a given side-chain, as clearly illustrated for Asn 26 and Phe 41 that reflects differences in $\Delta\omega$ values at each site. Notably, the intrinsic transverse relaxation rates of the H^{γ} protons of Glx residues, $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}} \rightarrow \infty)$, are relatively small ($<14\ s^{-1}$), as expected for side-chains that are primarily surface exposed and positions that are distal to the protein backbone. Values of

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178–3189

$R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}} \rightarrow \infty)$ become larger as the probe is closer to the backbone, as can be seen by a comparison of Asx and Glx profiles. The β protons of aromatic residues relax rapidly since these amino acids are localized primarily to hydrophobic cores and are immobilized through a myriad of contacts within the molecule.

We have also recorded a second set of dispersion profiles under different conditions, 10 °C, pH 6.6, H_2O, where the exchange parameters are varied ($p_1 = 2.1 \pm 0.05\%$, $k_{ex} = 320 \pm 10\ s^{-1}$). The chemical shift differences, $\Delta\varpi$ (ppm), are expected to be only very slightly dependent on the conditions of the experiment so that a strong correlation between $\Delta\varpi$ values recorded on the same system with different exchange parameters provides a cross-validation of the methodology. Figure 5

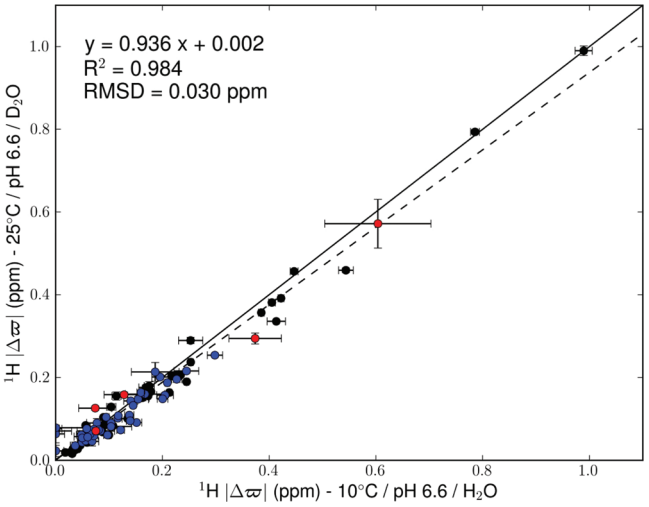


Figure 5. Linear correlation plot of $|\Delta\varpi|$(25 °C, pH 6.6, D_2O; *Y*-axis) vs $|\Delta\varpi|$(10 °C, pH 6.6, H_2O; *X*-axis) values. Values from aliphatic nonmethyl, methyl and aromatic side chains are color coded blue, black, and red, respectively. The line $y = x$ (solid) and the best fit line (dashed) are indicated, along with Pearson's correlation coefficient squared and the rmsd between the two data sets. Errors in $|\Delta\varpi|$ values are denoted by vertical and horizontal bars.

shows the linear correlation plot of $\Delta\varpi$(25 °C, pH 6.6, D_2O; *Y*-axis) and $\Delta\varpi$(10 °C, pH 6.6, H_2O; *X*-axis) values, color coded according to side-chain type (blue: aliphatic nonmethyl, black: aliphatic methyl, red: aromatic), along with the line $y = x$ and the best fit line (dashed). It is clear that an excellent correlation is obtained, providing a strong measure of confidence in the extracted chemical shifts. The small deviations that are observed have also been noted in comparisons of $\Delta\varpi$ values for ^{15}N and $^{1}H^{N}$ nuclei obtained via analysis of dispersion profiles recorded at 25 and 10 °C (unpublished data).

The CPMG relaxation dispersion spectra that have been recorded to obtain the aliphatic ^1H dispersion profiles discussed above can also be analyzed to produce the corresponding curves for methyl groups. Cross-peaks derived from methyls of Ala, Ile$^{\gamma 2}$, Leu, Met and Val are straightforward to fit; by contrast, the situation for Thr and Ile$^{\delta 1}$ is complicated by the fact that carbons adjacent to the methyl groups in question are not completely deuterated, although the predominant isotopomer is of the CD or CD_2 variety for Thr$^{\beta}$ and Ile $C^{\gamma 1}$, respectively. In spectra of Im7 recorded at 18.8 T we have found that resolution of the deuterium isotope shifted correlations is sufficient to separate Thr and Ile $C^{\delta 1}$ methyl peaks that are adjacent to $C^{\beta}D$ and $C^{\gamma 1}D_2$ spin systems, respectively, so that robust measures of exchange can be obtained at these methyl positions as well without complications from magnetization transfer due to $^1H-^1H$ homonuclear scalar couplings.

Figure 6 plots ^1H and ^{13}C dispersion profiles derived from a number of side chains in Im7 (25 °C, pH 6.6, D_2O) to illustrate the complementarity of experiments that focus on both ^1H and ^{13}C spins. This complementarity is to be expected, at least for methyl groups, where it is known that ^{13}C probes are very sensitive to changes in side-chain dihedral angles between exchanging ground and excited states via the γ-gauche effect,[60,61] while the ^1H spin senses 'external' structural perturbations, such as those due to changes in ring currents (see below). The results from the present study show that for some residues, for example Glu 2 or Asn 26, 'similar' dispersion profiles are obtained for both ^1H and ^{13}C side-chain nuclei. For others, however, (Leu 37, Glu 46, Val 42) there are large differences between ^1H and ^{13}C profiles. Not surprisingly, there are instances when ^1H dispersions are large, while ^{13}C profiles from neighboring probes are small and vice versa. On average, dispersions might be expected to be larger for side-chain carbons than protons, based on a comparative statistical analysis of ^{13}C and ^1H side-chain chemical shift distributions reported by the BMRB (http://www.bmrb.wisc.edu/ref_info) showing a larger standard deviation (in Hz) for carbon. However, changes in $\Delta\omega$ values in cases where the probe in question changes orientation/distance with respect to an aromatic ring may be significantly larger for ^1H than ^{13}C, both because of the factor of 4 increase in gyromagnetic ratio and the fact that ^1H spins are often closer to the aromatic(s) in question. This increases the value of the ^1H spin as a probe of structure.

The distribution of ^1H probes that are available for analysis in Im7 is illustrated in Figure 7A, where protons from 29 methyl groups (black), 13 Asx (H^{β}), 13 Glx (H^{γ}), 4 Ser (H^{β}) (all in blue), and 9 aromatic (H^{δ}, red) residues are shown. Two orientations of the crystal structure of the ground state of Im7[64] are illustrated in Figure 7B showing the overall structure, including the four helices that comprise the native state of the protein. In addition, a number of residues are highlighted that are discussed below. The great majority of the methyl groups and aromatic side chains are localized to the interior of the protein serving as reporters of changes in hydrophobic packing associated with the conformational transition probed by the CPMG relaxation dispersion experiment. In contrast, Asx and Glx (and the side chains of Tyr 55, Tyr 56, and Trp 75) are localized to the surface of the protein, so that structural changes to surface positions can be studied as well.

Although details of the structure of the excited state must await an in-depth characterization of chemical shifts and residual dipolar couplings, as we have done in the past in studies of the FF domain,[22,24] a qualitative inspection of the backbone chemical shifts along with the side-chain carbonyl shifts that we have reported previously[29] allows us to conclude with confidence that H3 (helix 3) is no longer folded in the excited state (red in Figure 7B). The measured relaxation dispersion data reported here are consistent with this. Figure 7C shows values of $|\Delta\varpi|$ color-coded onto the side-chain protons that are available for analysis in the present set of experiments. Significant changes in chemical shifts are observed for probes that are close to or part of H3. These include Leu 53 ($|\Delta\varpi^{\delta 1}| = 0.45$ ppm, $|\Delta\varpi^{\delta 2}| = 0.36$ ppm) and Ile 54 ($|\Delta\varpi^{\delta}| = 0.99$ ppm, $|\Delta\varpi^{\gamma 2}| = 0.79$ ppm) that are located in the middle of H3 in the native state. The large $|\Delta\varpi|$ values reflect the proximity of the methyl protons to His 47 (Leu 53) and Phe 41,Tyr 55 (Ile 54), leading

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

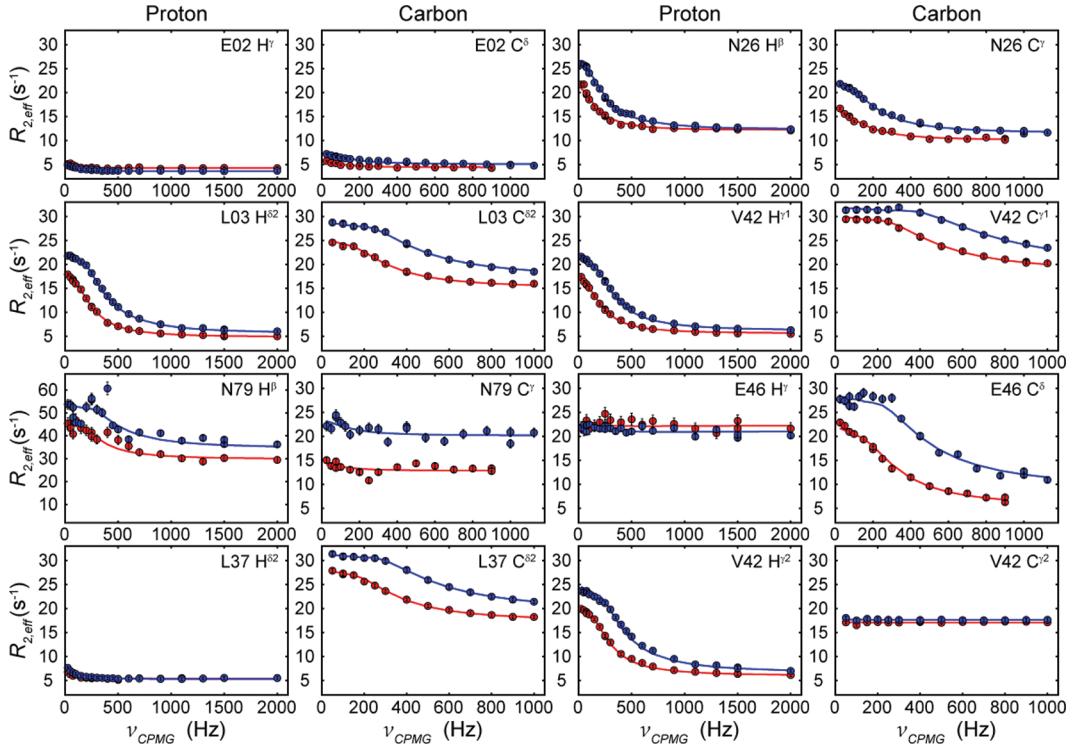


Figure 6. Representative CPMG relaxation dispersion profiles for select ^{13}C and ^{1}H nuclei from a given residue, indicated in the top right of each panel. All data were recorded on a sample of fractionally deuterated, U−^{13}C Im7, prepared as described in Materials and Methods at 11.7 T (red) and 18.8 T (blue). Side-chain carbonyl dispersions were recorded as described previously[29] while methyl ^{13}C dispersions were obtained using a sample produced with 1-^{13}C glucose.[62] Stereospecific assignment of the prochiral methyl groups were obtained by the method of Neri.[63] Carbon and proton dispersion data were analyzed separately.

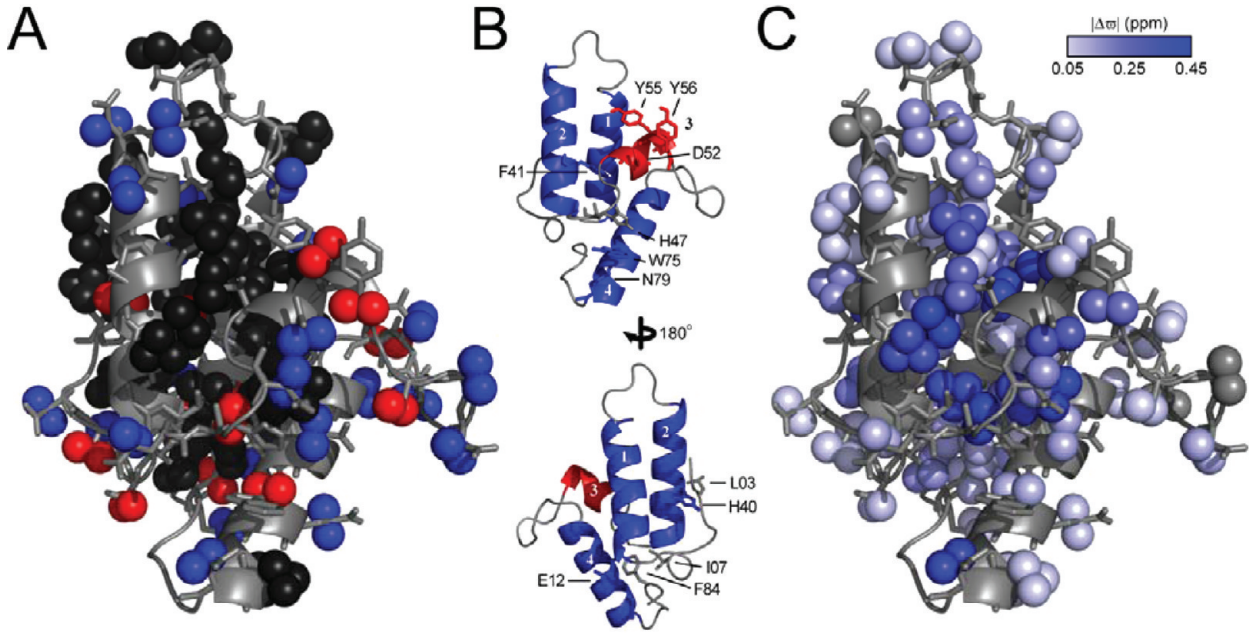


Figure 7. (A) Distribution of side-chain ^{1}H probes (balls) in Im7 that are available for analysis via ^{1}H CPMG relaxation dispersion. Twenty-nine methyl groups (black), 13 Asx (H^{β}), 13 Glx (H^{γ}), 4 Ser (H^{β}) (all in blue), and 9 aromatic (H^{β}, red) residues are indicated. (B) Two views of the crystal structure of the ground state of Im7,[64] highlighting the overall structure including H3 that is unfolded in the intermediate (red), and several residues that are discussed in the text. Also shown on the structure are key sites that have large $|\Delta\varpi|$ values. (C) Values of $|\Delta\varpi|$ obtained in the analysis of the ^{1}H dispersion experiment are color-coded onto the side-chain protons. In cases where values of $|\Delta\varpi|$ are not available the side-chain protons are colored gray.

to large upfield shifts in spectra of the native conformation (see Figure 3B). Val 42 ($|\Delta\varpi^{\gamma 1}|$ = 0.34 ppm, $|\Delta\varpi^{\gamma 2}|$ = 0.46 ppm) at the C-terminal end of H2 also shows large shift changes, consistent with the view that this helix is shorter in the excited state, a conclusion based on backbone chemical shifts that have been measured by other relaxation dispersion experiments.

dx.doi.org/10.1021/ja210711v | J. Am. Chem.Soc. 2012, 134, 3178−3189

Other residues such as Leu 03 ($|\Delta\varpi^{\delta 1}|$ = 0.18 ppm, $|\Delta\varpi^{\delta 2}|$ = 0.39 ppm), Thr 45 ($|\Delta\varpi^{\gamma}|$ = 0.38 ppm) and Leu 38 ($|\Delta\varpi^{\delta 1}|$ = 0.29 ppm, $|\Delta\varpi^{\delta 2}|$ = 0.05 ppm) also show substantial shift changes and notably in almost all cases the methyl groups in question are proximal to aromatic side-chains. It seems very likely, therefore, that ^{1}H $|\Delta\varpi|$ values will provide useful 'long-range' information connecting side-chains distant in primary sequence. Notably, small chemical shift changes can also provide structural insights, as in the case of Ile 7 of H1 ($|\Delta\varpi^{\delta 1}|$ = 0.13 ppm, $|\Delta\varpi^{\gamma 2}|$ = 0.15 ppm). Both δ1 and γ2 protons are significantly upfield shifted in the ground state (see Figure 3), due to the proximity of Phe 84 of H4, with only small changes in shifts in the excited conformation. This may indicate that helices 1 and 4 remain docked in the intermediate, consistent with the small changes in both the side-chain carbonyl and H^{β} shifts of Glu 12 that is known to form a salt-bridge in the ground state with Lys 73 of H4.

There are also substantial changes in proton shifts for a number of aromatic and side-chain Asx/Glx moieties in addition to the methyl groups mentioned above. Interestingly only a single ($^{13}C^{\beta}$,$^1H^{\beta}$) correlation is observed in spectra of His 47 ($|\Delta\varpi^{\beta}|$ = 0.57 ppm) while no correlations can be found for His 40, indicating that both of these residues are sensitive to the exchange process. His 40 is proximal to Leu 3, mentioned above, at the N-terminus of the protein suggesting that this region of Im7 may undergo a substantial rearrangement in the excited state. A large change in chemical shift is observed for one of the two β protons of Phe 41 (H2, $|\Delta\varpi^{\beta}|$ = 0.42, 0.13 ppm), that is centrally located in the protein. The large ^{1}H side-chain $|\Delta\varpi|$ values measured for other positions of H2 and the small backbone shift differences for the majority of this helix are consistent (unpublished data) with this region retaining helical structure in the excited state, with reshuffling of side-chain interactions. Not surprisingly changes in β-proton chemical shifts are also observed for Tyr 56 ($|\Delta\varpi^{\beta}|$ = 0.29 ppm, 0.16 ppm) and for Asp 52 ($|\Delta\varpi^{\beta}|$ = 0.48 ppm, 0.09 ppm), both located in H3 that is unfolded in the excited conformer. Finally, significant H^{β} shift changes are measured for Asn 79 ($|\Delta\varpi^{\beta}|$ = 0.57 ppm, 0.21 ppm). In the ground state its resonance position is upfield shifted (Figure 3A) due to the proximal Trp 75 ring. The side-chain of Asn 79, located at the C-terminal end of H4, forms a hydrogen bond with the backbone carbonyl of Trp 75 in the native conformation[64] and the shift change thus likely reflects fraying of this helix in the excited state.

CONCLUDING REMARKS

An approach has been described for the measurement of side-chain proton CPMG relaxation dispersion profiles for fractionally deuterated proteins generated via overexpression using {^{13}C,^{1}H}-glucose as the only carbon source in D_2O media. 'Isolated' ^{1}H spins are produced at the H^{β} positions of aromatic amino acids, Asx, Cys, and Ser, at H^{γ} sites of Glx and for most of the methyl groups. This provides for a large number of side-chains that can be probed by the relaxation dispersion technique, leading to the extraction of $|\Delta\varpi|$ values to be subsequently used in calculations of the structure of the excited state. Particularly important in this regard is the sensitivity of ^{1}H chemical shifts to positions of aromatic residues that is expected to significantly improve the structures obtained. The experiment presented significantly extends previous methodology, largely focused on measurements of backbone chemical shift differences. As such it adds to a growing list of experiments that 'bring into focus' excited protein states that have remained recalcitrant to study using the traditional tools of structural biology.

ASSOCIATED CONTENT

Supporting Information

Descriptions of the growth conditions used to produce the fractionally ^{2}H Im7 sample and of the experimental approach for quantifying the extent of deuteration/protonation at key side-chain positions in proteins using {^{13}C,^{1}H}-glucose as the only carbon source in D_2O media. Discussion of the filtering schemes of Figure 2 used to select against $^{13}CH_2$/$^{13}CH_2D$ isotopomers. Full author list for reference 18. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

kay@pound.med.utoronto.ca

ACKNOWLEDGMENTS

The authors thank Dr. Sara Whittaker and Professors Geoff Moore and Sheena Radford for the gift of Im7 constructs and Dr. Ranjith Muhandiram for assistance with the NMR experiments. A.L.H. acknowledges the National Science Foundation for postdoctoral support (OISE-0852964). P.L. is supported by the Swedish Research Council. This work was funded by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada (L.E.K.). L.E.K. holds a Canada Research Chair in Biochemistry.

REFERENCES

(1) Carr, H. Y.; Purcell, E. M. *Phys. Rev.* **1954**, *54*, 630−638.
(2) Meiboom, S.; Gill, D. *Rev. Sci. Instrum.* **1958**, *29*, 688−691.
(3) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *339*, 204−238.
(4) Korzhnev, D. M.; Kay, L. E. *Acc. Chem. Res.* **2008**, *41*, 442−451.
(5) Skrynnikov, N. R.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 12352−12360.
(6) Auer, R.; Neudecker, P.; Muhandiram, D. R.; Lundstrom, P.; Hansen, D. F.; Konrat, R.; Kay, L. E. *J. Am. Chem. Soc.* **2009**, *131*, 10832−10833.
(7) Loria, J. P.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1999**, *121*, 2331−2332.
(8) Tollinger, M.; Skrynnikov, N. R.; Mulder, F. A. A.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 11341−11352.
(9) Ishima, R.; Torchia, D. *J. Biomol. NMR* **2003**, *25*, 243−248.
(10) Hansen, D. F.; Vallurupalli, P.; Lundstrom, P.; Neudecker, P.; Kay, L. E. *J. Am. Chem. Soc.* **2008**, *130*, 2667−2675.
(11) Ishima, R.; Baber, J.; Louis, J. M.; Torchia, D. A. *J. Biomol. NMR* **2004**, *29*, 187−98.
(12) Lundstrom, P.; Hansen, D. F.; Kay, L. E. *J. Biomol. NMR* **2008**, *42*, 35−47.
(13) Lundstrom, P.; Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J. Am. Chem. Soc.* **2009**, *131*, 1915−1926.
(14) Vallurupalli, P.; Hansen, D. F.; Stollar, E. J.; Meirovitch, E.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18473−18477.
(15) Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J. Am. Chem. Soc.* **2008**, *130*, 8397−8405.
(16) Igumenova, T. I.; Brath, U.; Akke, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2007**, *129*, 13396−13397.
(17) Vallurupalli, P.; Hansen, D. F.; Kay, L. E. *J. Am. Chem. Soc.* **2008**, *130*, 2734−2735.
(18) Shen, Y.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685−4690.
(19) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615−9620.

(20) Wishart, D. S.; Arndt, D.; Berjanskii, M.; Tang, P.; Zhou, J.; Lin, G. *Nucleic Acids Res.* **2008**, *36*, W496−502.
(21) Vallurupalli, P.; Hansen, D. F.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11766−11771.
(22) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312−1316.
(23) Bouvignies, G.; Vallurupalli, P.; Hansen, D. F.; Correia, B. E.; Lange, O.; Bah, A.; Vernon, R. M.; Dahlquist, F. W.; Baker, D.; Kay, L. E. *Nature* **2011**, *477*, 111−114.
(24) Korzhnev, D. M.; Vernon, R. M.; Religa, T. L.; Hansen, A. L.; Baker, D.; Fersht, A. R.; Kay, L. E. *J. Am. Chem. Soc.* **2011**, *133*, 10974−10982.
(25) Lundstrom, P.; Kay, L. E. *J. Biomol. NMR* **2009**, *44*, 139−155.
(26) Lundstrom, P.; Vallurupalli, P.; Religa, T. L.; Dahlquist, F. W.; Kay, L. E. *J. Biomol. NMR* **2007**, *38*, 79−88.
(27) Otten, R.; Villali, J.; Kern, D.; Mulder, F. A. *J. Am. Chem. Soc.* **2010**, *132*, 17004−17014.
(28) Baldwin, A. J.; Religa, T. L.; Hansen, D. F.; Bouvignies, G.; Kay, L. E. *J. Am. Chem. Soc.* **2010**, *132*, 10992−5.
(29) Hansen, A. L.; Kay, L. E. *J. Biomol. NMR* **2011**, *50*, 347−355.
(30) Paquin, R.; Ferrage, F.; Mulder, F. A.; Akke, M.; Bodenhausen, G. *J. Am. Chem. Soc.* **2008**, *130*, 15805−7.
(31) Mulder, F. A.; Akke, M. *Magn. Reson. Chem.* **2003**, *41*, 853−865.
(32) Mulder, F. A. A.; Skrynnikov, N. R.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 967−975.
(33) Esadze, A.; Li, D. W.; Wang, T.; Bruschweiler, R.; Iwahara, J. *J. Am. Chem. Soc.* **2011**, *133*, 909−919.
(34) Wishart, D. S.; Case, D. A. *Methods Enzymol.* **2001**, *338*, 3−34.
(35) Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Mei Ono, A.; Guntert, P. *Nature* **2006**, *440*, 52−57.
(36) Shekhtman, A.; Ghose, R.; Goger, M.; Cowburn, D. *FEBS Lett.* **2002**, *524*, 177−12.
(37) Guo, C.; Tugarinov, V. *J. Biomol. NMR* **2009**, *46*, 127−133.
(38) Gerstein, M. *J. Mol. Biol.* **1997**, *274*, 562−576.
(39) Le Duff, C. S.; Whittaker, S. B.; Radford, S. E.; Moore, G. R. *J. Mol. Biol.* **2006**, *364*, 824−85.
(40) Levitt, M.; Freeman, R. *J. Magn. Reson.* **1978**, *33*, 473−476.
(41) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335−338.
(42) Geen, H.; Freeman, R. *J. Magn. Reson.* **1991**, *93*, 93−141.
(43) Shaka, A. J.; Barker, P. B.; Freeman, R. *J. Magn. Reson.* **1985**, *64*, 547−552.
(44) Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663−10665.
(45) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393−399.
(46) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(47) Kneller, D.; Kuntz, I. *J. Cell. Biochem.* **1993**, No. Suppl. 17C, 254.
(48) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*; Cambridge University Press: Cambridge, 1988.
(49) Ishima, R.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Torchia, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 10534−10542.
(50) Bothner-By, A. A.; Stephens, R. L.; Lee, J. M.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.
(51) Stryer, L. *Biochemistry*; fourth ed.; W.H. Freeman and Company: New York, 1995.
(52) Otten, R.; Chu, B.; Krewulak, K. D.; Vogel, H. J.; Mulder, F. A. *J. Am. Chem. Soc.* **2010**, *132*, 2952−60.
(53) Cleland, W. W. *Arch. Biochem. Biophys.* **2005**, *433*, 2−12.
(54) Dubrow, R.; Pizer, L. I. *J. Biol. Chem.* **1977**, *252*, 1539−1551.
(55) Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J. Phys. Chem.* **2008**, *112*, 5898−5904.
(56) Mittermaier, A.; Kay, L. E. *J. Biomol. NMR* **2002**, *23*, 35−45.
(57) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 185−204.
(58) Capaldi, A. P.; Kleanthous, C.; Radford, S. E. *Nat. Struct. Biol.* **2002**, *9*, 209−216.
(59) Capaldi, A. P.; Shastry, M. C.; Kleanthous, C.; Roder, H.; Radford, S. E. *Nat. Struct. Biol.* **2001**, *8*, 68−72.
(60) Mulder, F. A. *ChemBioChem* **2009**, *10*, 1477−1479.
(61) Hansen, D. F.; Neudecker, P.; Vallurupalli, P.; Mulder, F. A.; Kay, L. E. *J. Am. Chem. Soc.* **2010**, *132*, 42−43.
(62) Lundstrom, P.; Teilum, K.; Carstensen, T.; Bezsonova, I.; Wiesner, S.; Hansen, D. F.; Religa, T. L.; Akke, M.; Kay, L. E. *J. Biomol. NMR* **2007**, *38*, 199−212.
(63) Neri, D.; Szyperski, T.; Otting, G.; Senn, H.; Wüthrich, K. *Biochemistry* **1989**, *28*, 7510−7516.
(64) Dennis, C. A.; Videler, H.; Pauptit, R. A.; Wallis, R.; James, R.; Moore, G. R.; Kleanthous, C. *Biochem. J.* **1998**, *333* (Pt 1), 183−191.