Accelerated Publications

Quantitative ^{13}C and ^{2}H NMR Relaxation Studies of the 723-Residue Enzyme Malate Synthase G Reveal a Dynamic Binding Interface†

Vitali Tugarinov and Lewis E. Kay*

Departments of Medical Genetics, Biochemistry, and Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A8

Received September 29, 2005; Revised Manuscript Received November 1, 2005

ABSTRACT: A detailed understanding of molecular recognition is predicated not only on high-resolution static structures of the free and bound states but also on information about how these structures change with time, that is, molecular dynamics. Here we present a deuterium (^{2}H) and carbon (^{13}C) NMR relaxation study of methyl side chain dynamics in the 82 kDa enzyme malate synthase G (MSG) that is a promising target for the development of new antibiotic agents. It is shown that excellent agreement between ^{2}H- and ^{13}C-derived measures of dynamics is obtained, with correlation coefficients exceeding 0.95. The binding interface formed by MSG and its substrates is found to be highly dynamic in the ligand-free state of the enzyme with rigidification upon binding substrate. This study establishes that detailed, quantitative information about methyl side chain dynamics can be obtained by NMR on proteins with molecular masses on the order of 100 kDa and opens up the possibilities for studies of motion in a large number of important systems.

It is widely accepted that biomolecular recognition depends on the structural properties of the interacting molecules. Historically, much of our knowledge of molecular structure has been derived from static pictures of proteins, using either X-ray or NMR analyses. Such structures often provide detailed models that in some cases can explain specificity and function. However, despite recent advances in proteomics and the accelerated pace of high-throughput protein structure determination, our understanding of the mechanism and driving forces that underlie biomolecular interactions is, in many cases, rudimentary (*1*, *2*).

One of the major difficulties is that molecules are not static and often their dynamics are critical in recognition processes or for function. Characterizing the motions on a site-specific basis and relating such dynamics to function or stability in a quantitative way remain major challenges. In principle, NMR spectroscopy is a valuable probe of motion over a wide range of time scales, and in the past decade, a large number of experimental approaches that measure dynamics at different positions along both the backbone and side chains of proteins (*3*−*5*), in particular, have been derived. A limitation of the methodology has been, however, that applications are restricted to relatively small systems, typically with molecular masses of less than ∼40 kDa; many interesting and important molecules lie outside this window. Recently, the development of new labeling approaches and a new class of NMR experiment that relies on transverse relaxation-optimized spectroscopy (TROSY)[1] for backbone amide (*6*), methyl (*7*), and methylene (*8*, *9*) groups, as well as significant sensitivity improvements in instrumentation, has led to increases in the size of proteins that can be studied by NMR.

One such large protein that has been studied in considerable detail by multidimensional NMR methods in our laboratory is the 82 kDa single-polypeptide chain malate synthase G (MSG, 723 residues). MSG and isocitrate lyase are the two enzymes of the "glyoxylate shunt", a biochemical bypass used by microorganisms (bacteria, yeast, and fungi) for biosynthesis under anaerobic conditions (*10*). Enzymes of the glyoxylate bypass have been implicated as virulence factors in several pathogens, including *Mycobacterium tuberculosis* (*11*−*13*). The importance of these glyoxylate shunt enzymes is underscored by the fact that tuberculosis is estimated to exist as a dormant infection in more than 2 billion people and to result in more than 1.7 million deaths annually (*14*). Because the glyoxylate bypass is not found in humans, MSG represents a promising target for the development of new antibiotic agents. As a result, there have been a number of structural studies of this protein. The 1.95 Å resolution X-ray structure of the ternary abortive complex of MSG with pyruvate (glyoxylate mimic) and acetyl-CoA has been determined recently (*15*), showing that the enzyme

† This work was supported by a grant from the Canadian Institutes of Health Research (CIHR) to L.E.K. V.T. acknowledges the financial support of the CIHR in the form of a postdoctoral fellowship. L.E.K. holds a Canada Research Chair in Biochemistry.

* To whom correspondence should be addressed. E-mail: kay@pound.med.utoronto.ca. Fax: (416) 978-6885. Telephone: (416) 978-0741.

[1] Abbreviations: MSG, malate synthase G; CSA, chemical shift anisotropy; TROSY, transverse relaxation-optimized spectroscopy; HMQC, heteronuclear multiple-quantum coherence; HSQC, heteronuclear single-quantum coherence; MD, molecular dynamics; ILV, isoleucine, leucine, valine; TIM, triosephosphate isomerase; rmsd, root-mean-square deviation.

10.1021/bi0519809 CCC: $30.25
Published on Web 11/16/2005

adopts essentially the same structure as in the complex with glyoxylate reported earlier (*16*). Recently, using new NMR approaches, we have determined the global fold of the apo form of MSG from NMR data exclusively (*17*), showing that the relative orientation of domains in the molecule does not change upon ligation (*17*, *18*).

We report here a methyl side chain dynamics study of Ile, Leu, and Val residues in MSG using both ^{13}C and ^{2}H spin relaxation probes of motion in apo- and pyruvate-acetyl-CoA-saturated forms of the enzyme that significantly complements the existing body of static structural information about this important molecule. In MSG, there are close to 300 Ile/Leu/Val methyl groups, with excellent coverage over the complete structure of the protein. Twenty-two of these methyl carbons are within 10 Å of one or more atoms of either pyruvate or acetyl-CoA and thus serve as excellent reporters on the dynamics in the vicinity of the active site. The measurements show that there is a remarkable level of structural plasticity in much but not all of this region in the apo structure that, for the most part, rigidifies upon ligand binding. This study shows for the first time that site-specific and quantitative dynamics information can be obtained for Ile/Leu/Val methyl groups in proteins with molecular masses on the order of 100 kDa and sets the stage for further dynamics studies of large proteins.

MATERIALS AND METHODS

NMR Sample. A U-[^{15}N,^{2}H], Ileδ1-[$^{13}CHD_2$]-labeled sample of MSG, which was available at the outset of the study (*8*, *9*), was used to record ^{2}H and ^{13}C relaxation rates (of Ile methyl groups) in the apo form of the protein. A second MSG sample that was U-[^{15}N,^{2}H], Ileδ1-[$^{13}CHD_2$], Leu,Val-[$^{13}CHD_2$,$^{13}CHD_2$]-labeled subsequently became available and was used for all remaining measurements, including measurement of both ^{2}H and ^{13}C Leu/Val methyl spin relaxation properties in apo-MSG and all Ile/Leu/Val methyl measurements in the bound state. Details of sample production are as described previously (*19*, *20*), with selective [$^{13}CHD_2$] labeling of methyl groups of Ile, Leu, and Val in an otherwise deuterated protein achieved by the addition of 60 mg of 2-keto-3-d_2-4-^{13}C,d_2-butyrate and 100 mg of 2-keto-3-methyl-^{13}C,d_2-3-d_1-4-^{13}C,d_2-butyrate to 1 L of D_2O-based minimal medium following the method of Gardner and Kay (*21*) and Goto et al. (*22*). Ileδ1-[$^{13}CHD_2$]-labeled and Ileδ1-[$^{13}CHD_2$], Leu,Val-[$^{13}CHD_2$,$^{13}CHD_2$]-labeled samples were at concentrations of 0.45 and 0.80 mM in protein, respectively, in 99% D_2O and contained 25 mM sodium phosphate buffer (pH 7.1, uncorrected), 20 mM $MgCl_2$, 5 mM DTT, and 0.05% NaN_3.

Assignments of Methyl Groups in the Ternary Abortive MSG−Pyruvate−Acetyl-CoA Complex. Ile, Leu, and Val methyl assignments (~95% completeness) for the apo form of MSG have been reported (*20*, *23*). Assignments of the corresponding methyl groups in the ternary complex (pyruvate−acetyl-CoA−MSG) have been obtained from a series of titrations in which first sodium pyruvate was added until no further changes in peak positions were noted (saturation), followed by the addition of acetyl-CoA until saturation. Since the binding of both ligands occurs in the intermediate to fast exchange regime (*18*), it is possible to follow the "trajectories" of cross-peak positions in two-dimensional spectra during both titration series for ~75% of the methyl resonances. The final concentrations of pyruvate and acetyl-CoA in the ternary, ILV-[$^{13}CHD_2$]-labeled MSG complex were 20 and 8 mM, respectively, reflecting the differences in affinities of the ligands for protein (1 mM and 270 μM, respectively). All NMR spectra were analyzed using *NMRView* (*24*).

^{2}H and ^{13}C Relaxation Measurements. All relaxation measurements were performed at 37 °C on a Varian Inova 600 MHz spectrometer equipped with a cryogenically cooled probe. ^{2}H transverse relaxation measurements in $^{13}CHD_2$ methyl isotopomers were performed using the scheme of Figure 6 of Tugarinov et al. (*9*). $R_{1\rho}$ and R_1 relaxation rates of ^{13}C spins in $^{13}CHD_2$ methyl groups have been measured using pulse schemes (given in the Supporting Information along with further experimental details) similar to those used by Torchia and co-workers for ^{13}C relaxation studies (*25*).

Data Analysis. Measured ^{2}H and ^{13}C relaxation rates are related to molecular dynamics because they depend on linear combinations of the power spectrum of the motion at a finite number of frequencies (for ^{2}H at 0, ω_D and $2\omega_D$ where ω_D is the ^{2}H Larmor frequency, for example). The power spectrum or spectral density can be most simply related to motional parameters using the model-free formalism (*26*, *27*) in which values of S^2_{axis} and τ_f are extracted per methyl site; S^2_{axis} values (order parameter squared) provide a measure of the amplitude of motion of the methyl 3-fold axis, with the time scale given by τ_f. ^{2}H and ^{13}C rates have been analyzed independently, and the extracted order parameters from the two very different methods were compared to cross-validate the methodology. This is particularly important given the very high molecular mass of MSG in relation to other proteins that have been studied in this manner to date (well over a factor of 2 larger), the fact that the ^{2}H methods are new (*9*), and the difficulties that have plagued ^{13}C spin relaxation studies of methyl groups in the past (*28*). Details of the fitting procedure, including the relevant equations for the relaxation rates and the form of the spectral density function, are given in the Supporting Information.

One of the powerful features of ^{2}H spin relaxation as a probe of molecular dynamics is that the relaxation of the deuteron is governed almost exclusively by the quadrupolar interaction that is well understood (*29*) and for which the relevant parameters, such as the quadrupolar coupling constant, have been quantified (*30*). In contrast, the interpretation of ^{13}C spin relaxation rates in $^{13}CHD_2$ methyl groups is more complicated because they depend on intra-methyl $^{1}H-^{13}C$ and $^{2}H-^{13}C$ dipolar spin interactions and methyl ^{13}C chemical shift anisotropy (CSA), as well as on dipolar contributions from both ^{1}H and ^{2}H spins that are external to the methyl group in question (*31*). Values for the assumed axially symmetric CSA have been measured for 14 Ile, 21 Leu, and 15 Val methyl groups in MSG using F_1-coupled $^{1}H-^{13}C$ HSQC spectra of ILV-[$^{13}CH_3$]-labeled MSG, as described previously (*32*), along with ^{2}H-derived (CSA-independent) S^2_{axis} values. The obtained average values (standard deviations) are 17.8 (6.2), 30.0 (5.1), 25.2 (6.3) ppm for Ile, Leu, and Val, respectively, and these values have been used in all analyses.

Computations for [^{2}H],ILV-[$^{13}CHD_2$]-labeled MSG show that contributions to ^{13}C relaxation rates from dipolar

interactions with external ^{2}H nuclei can be assigned by placing a single deuteron between 1.8 and 1.9 Å from the methyl carbon in question. An average distance of 1.8 Å was used uniformly. Contributions from dipolar fields due to external protons were estimated on a residue-specific basis by calculating the effective distance $(\sum_{ext} r_{CHext}^{-6})^{-1/6}$ for each methyl of interest (distance between the methyl carbon and all external protons) and then using this value in the course of the data fitting (see the Supporting Information). We have also assumed random protonation at the level of 3% (derived from the residual proton content of [^{2}H]-D-glucose, for example) in the analysis. The calculations establish that the contributions from external protons can be modeled as being derived from a single proton 3.4 (Ile), 2.6 (Leu), or 2.7 Å (Val) from the methyl carbon, on average. Note that the smaller distances for Leu and Val reflect the fact that both methyls of the same residue are $^{13}CHD_2$ (i.e., that there is a proximal proton to each methyl) while only one of the two methyls in Ile is protonated ($^{13}CHD_2$) in our labeling scheme.

RESULTS AND DISCUSSION

NMR Studies of Side Chain Methyl Dynamics in High-Molecular Mass Proteins. Methyl groups have long been used in the study of protein dynamics because their NMR spectra are usually of high sensitivity and resolution, and because they report on motions in protein hydrophobic cores, which are often important for function (*33*−*35*). In addition, methyl groups are usually well dispersed throughout the protein primary sequence (*36*) so that excellent coverage of the dynamics across the protein can be obtained. Here we have studied methyl groups of Ile, Leu, and Val as reporters of motion; in a survey of more than 1000 unrelated proteins, these three residues comprise 21% of the database (*37*) and thus are generally well represented (22% in MSG). We have chosen to prepare U-[^{15}N,^{2}H], Ileδ1-[$^{13}CHD_2$],Leu,Val-[$^{13}CHD_2$,$^{13}CHD_2$]-labeled protein using appropriately labeled α-keto acid compounds as biosynthetic precursors during protein production (*38*) (see Materials and Methods) so that both ^{2}H and ^{13}C spin relaxation experiments can be performed on the same sample. Selected regions of the two-dimensional ^{1}H−^{13}C correlation maps of ILV-[$^{13}CHD_2$] samples of apo- and pyruvate-acetyl-CoA-MSG are shown in Figure 1a,b. It is clear that only a single methyl isotopomer is present (essentially 100% incorporation), critical in applications to high-molecular mass proteins where both resolution and sensitivity are concerns. Recently, Homans and co-workers used a similar isotope labeling scheme in ^{13}C methyl relaxation studies, although limited to Val γ-methyl positions (*39*).

To date, applications of methyl spin relaxation have been restricted to studies of small proteins and existing ^{2}H spin relaxation methods, in particular, were quite limiting in terms of molecular size (less than ∼30 kDa). With this in mind, a suite of $^{13}CH_2D$-TROSY-based ^{2}H relaxation experiments were recently developed that are suitable for much larger systems (*9*). In addition, for large proteins (tumbling times in excess of ∼20 ns), it is also possible to obtain accurate values of S^2_{axis} by recording the decay of transverse ^{2}H relaxation in $^{13}CHD_2$-labeled methyl groups (*9*). Irrespective of whether $^{13}CH_2D$ or $^{13}CHD_2$ probes are employed, the attractive feature of ^{2}H measurements relates to the fact that

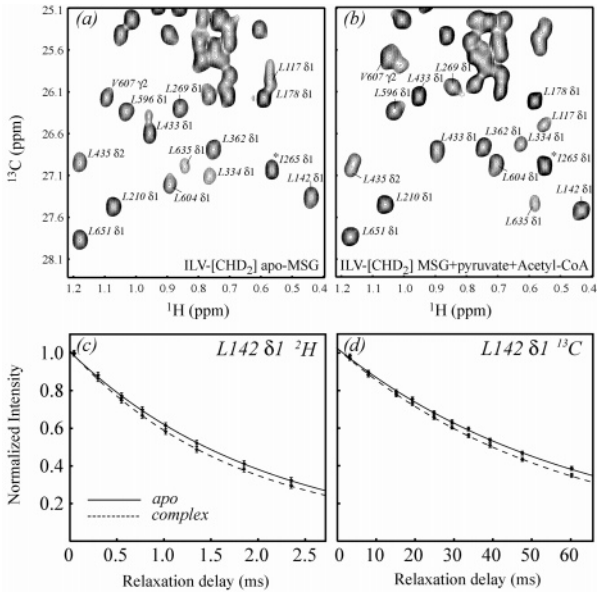

FIGURE 1: Selected region of the two-dimensional ^{1}H−^{13}C spectrum (600 MHz) corresponding to the first point of ^{13}C $R_{1\rho}$ measurements for (a) the apo form of ILV-[$^{13}CHD_2$]-labeled MSG and (b) the ternary complex of ILV-[$^{13}CHD_2$]-labeled MSG with pyruvate and acetyl-CoA. Separated peaks are labeled with their corresponding assignments in both protein states. Peaks of Ileδ1 methyls folded in the ^{13}C dimension are labeled with asterisks. The intensity decay curves fit to monoexponential functions are shown for the L142δ1 methyl group for ^{2}H $R(CHD_2)$ measurements (c) and ^{13}C $R_{1\rho}$ measurements (d). Exponential fits for data obtained for apo-MSG and the pyruvate−acetyl-CoA−saturated MSG complex are shown with solid and dashed lines, respectively.

the relaxation properties can be interpreted to excellent approximation by consideration of only the ^{2}H quadrupolar interaction (*40*). In contrast, the interpretation of ^{13}C relaxation in $^{13}CH_3$ methyl groups is complicated by interference effects between the dipolar fields created by the three protons, leading to nonexponential decay of the signal (*28*, *41*). In this regard, the use of a $^{13}CHD_2$ isotopomer is particularly appealing for ^{13}C studies since these interference effects are eliminated. However, even here the relaxation is more complex than for ^{2}H. For example, using an average set of motional parameters for Ile (Leu/Val) residues [S^2_{axis} = 0.6 (0.7); τ_f = 20 ps (40 ps); ^{13}C CSA = 18 ppm (25 ppm)] obtained from the [^{2}H],ILV-[$^{13}CHD_2$]-labeled MSG sample used in this work, the contribution of the main ^{13}C−^{1}H dipolar interaction to the cumulative ^{13}C relaxation rate is 78% (72%) for Ile (Leu/Val) methyl groups. Additional contributions include (i) intra-methyl ^{13}C−^{2}H dipolar effects comprising 10% (9%); (ii) CSA contributions of 5% (9%) at 600 MHz; (iii) dipolar interactions with external ^{1}H nuclei (i.e., outside the methyl of interest), 2% (6%); and (iv) dipolar contributions from interactions with external deuterons, 5% (4%). It is especially difficult to account accurately for contributions from CSA and external nuclei because of their variability from site to site; however, the influence from external protons can be minimized by high levels of sample deuteration.

The increased complexity in the interpretation of ^{13}C relaxation data in terms of motional parameters relative to ^{2}H measurements is offset to some extent by the sensitivity gains in the ^{13}C-based experiments. On average, the ^{13}C experiments are 3.3-fold more sensitive for MSG. This

reflects the fact that even in the absence of relaxation losses the transfer efficiency from ^{13}C to ^{2}H and back in a $^{13}CHD_2$ group, which is a prerequisite for the ^{2}H experiments (*9*), is 0.5. The $^{13}CHD_2$ labeling scheme employed in this study affords an opportunity to assess both ^{2}H and ^{13}C relaxation on a single sample and within the same isotopomer, and hence assess whether the contributions to the ^{13}C relaxation detailed above compromise the extraction of robust motional parameters. A previous comparison of ^{2}H- and ^{13}C-derived measures of order at methyl ($^{13}CHD_2$) positions for the HIV-1 protease would suggest not (*31*); however, in that study, slightly elevated S^2_{axis} values were extracted from the ^{13}C relaxation data, likely resulting from artifacts introduced by rapid ^{1}H pulsing during the ^{13}C spin lock period (*42*).

Correlation between ^{13}C and ^{2}H Measures of Dynamics. Panels c and d of Figure 1 show typical decay curves of ^{2}H and ^{13}C transverse magnetization from ILV-[$^{13}CHD_2$]-labeled MSG that form the primary data for extraction of motional parameters. The number of methyl sites whose relaxation rates could be reliably quantified is somewhat lower than the total number of ILV sites in MSG [approximately 65% of the 44 Ile (δ1), 92 Val, and 140 Leu methyls could be measured in the apo form and 50% in the ligand-bound state] primarily because of cross-peak overlap in the two-dimensional relaxation spectra and, in the case of the ternary complex, incomplete assignments (see Materials and Methods). The correlations between ^{2}H-derived and ^{13}C-derived S^2_{axis} values in the apo (a−c) and ligated (d−f) states of the enzyme are shown in Figure 2, separated according to amino acid type. The data are highly correlated, with correlation coefficients ranging from 0.95 to 0.99 and root-mean-square differences in S^2_{axis} of 0.05−0.07. A statistically significant correlation is also observed between the measured order parameters and the inverse crystallographic temperature factors [PDB entry 1d8c (*16*) (Pearson's *R* of 0.4−0.6, depending on what residue types are considered)].

The results summarized in Figure 2 make it clear that despite the increased complexity in the interpretation of the ^{13}C relaxation data, robust measures of dynamics can be obtained (*31*). It is noteworthy that the level of agreement between $S^2_{axis}(^{2}H)$ and $S^2_{axis}(^{13}C)$ is best for Ile where the ^{13}C CSA is smallest. The rms differences between S^2_{axis} values were somewhat worse when data obtained at 800 MHz were analyzed, likely reflecting the substantially higher contributions of CSA relaxation at this field (a factor of approximately 2 increase in the contribution to the total rate relative to 600 MHz) and the site-specific variation in CSA values that is not taken into account. A second source of error in the interpretation of the ^{13}C data may result from the assumption of fixed methyl geometry. Analysis of a combined set of $^{1}D_{C-H}$ and $^{1}D_{C-C}$ dipolar couplings from the studies of Ottiger and Bax (*43*) and Mittermaier and Kay (*44*) showed that $^{1}D_{C-H}/^{1}D_{C-C}$ ratios, which relate directly to C−H bond length as well as to the angle between the C−H bond and the methyl 3-fold rotation axis, are subject to somewhat higher variability in Val and Leu methyls than the same ratios in Ile. Changes in methyl geometry could of course also affect the extracted ^{2}H-based measures of order (*30*). Finally, the majority of (primarily ^{13}C-derived) S^2_{axis} values of Val and Leu residues that exceed the theoretical upper limit of 1.0 were derived from R_2 relaxation rates that

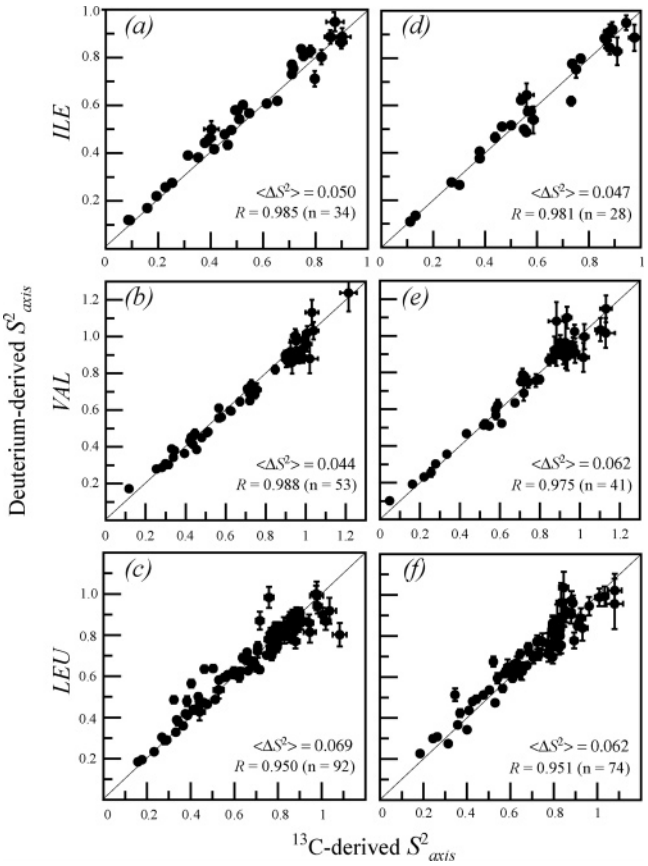

FIGURE 2: Linear correlation plots of ^{2}H-derived S^2_{axis} (*y*-axis) vs ^{13}C-derived S^2_{axis} (*x*-axis) for ILV methyls in (a−c) apo-MSG and (d−f) the pyruvate−acetyl-CoA−MSG abortive ternary complex. Average rmsd values obtained between ^{13}C- and ^{2}H-derived S^2_{axis} distributions, $\langle \Delta S^2 \rangle$, are indicated along with Pearson correlation coefficients, *R*. A diagonal line ($y = x$) is included in each plot.

showed statistically significant increases as a function of decreasing ^{13}C spin lock fields (from 2.0 to 1.4 kHz), indicating that chemical exchange processes may not be completely quenched at the highest spin lock field employed (see below). On average, however, the good agreement between $S^2_{axis}(^{13}C)$ and $S^2_{axis}(^{2}H)$ (Figure 2) implies that contributions from chemical exchange are small. This is further substantiated by the fact that the average difference between $S^2_{axis}(^{13}C$; 2.0 kHz) and $S^2_{axis}(^{13}C$; 1.4 kHz) is 0.003 $\pm$ 0.036 with an average of 0.071 $\pm$ 0.031 for the 10 largest differences.

Global Dynamic Properties. Panels a−c of Figure 3 show histograms of the ^{2}H-derived S^2_{axis} values obtained for ILV methyls of the apo form of MSG. Recent results from molecular dynamics (MD) simulations of Best et al. (*45*) indicate that distributions of S^2_{axis} values can be understood in terms of the rotameric preferences in methyl-containing side chains. For Ile and Leu residues, high, medium, and low values of S^2_{axis} reflect whether neither, one, or two of the χ_1 and χ_2 angles undergo rotameric transitions between wells, and give rise to a three-pronged distribution (Figure 3a,c). The two-pronged profile for Val (Figure 3b) results from distributions of residues that either are confined to a single χ_1 state or that undergo jumps between different states. It is noteworthy that although the positions of the peaks in the distributions of S^2_{axis} values for both Ile and Val residues in MSG correlate well with those expected both from MD

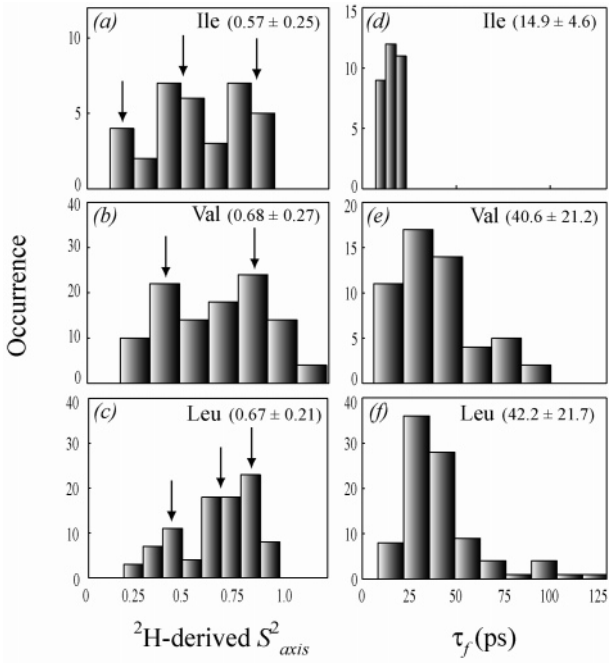

FIGURE 3: Histograms of S^2_{axis} values derived from ^{2}H relaxation measurements in ILV-[^{13}CHD$_2$]-labeled MSG for (a) Ile, (b) Val, and (c) Leu methyls. Arrows indicate approximate positions of maxima. Histograms of methyl internal rotation correlation times, τ_f (picoseconds), derived from ^{13}C relaxation data for (d) Ile, (e) Val, and (f) Leu methyls. Mean values together with standard deviations are indicated at the top of each distribution plot.

simulations and from experimental results on other proteins (*46*), the positions of the two peaks corresponding to low and intermediate S^2_{axis} values for Leu residues in MSG are shifted to somewhat higher values, reflecting an unusually large number of ordered Leu side chains in the large hydrophobic core of the protein.

The spin relaxation measures of dynamics presented here correlate well with other probes of motion. We have previously measured $^3J_{C\gamma C'}$ and $^3J_{C\gamma N}$ scalar couplings in Val residues as a prelude for stereospecific assignments of the prochiral methyls (*23*). Of 14 residues predicted from *J* couplings to be disordered or to have noncanonical side chain conformations, 10 have S^2_{axis} values of <0.5 for both methyls. Likewise, nine of the 14 Val side chains with motions on the micro- to millisecond time scale (*23*, *47*) have S^2_{axis} values of <0.5.

The distributions of local motional correlation times, τ_f, obtained by fitting ^{13}C R_1 and R_2 relaxation data for ILV residues in apo-MSG are shown in Figure 3d−f. As discussed previously, τ_f values reflect in a complex way the combined time scales of all motional processes that influence the relaxation, including internal methyl 3-fold rotation (*40*). The distribution of values for Ile is much narrower than for Leu or Val, with $\tau_{f,avg}$ smaller by a factor of 2.5, which likely reflects faster internal methyl rotation. Interestingly, a solid-state NMR study of ^{1}H R_1 methyl relaxation in compounds containing unbranched hydrocarbon chains (i.e., similar to Ileδ1 sites) and isopropyl groups (similar to Leu/Val) (*48*) showed similar methyl rotation barriers (11 and 12 kJ/mol, respectively).

Figure 4a displays the structure of MSG with the ILV methyls depicted as spheres color-coded according to their

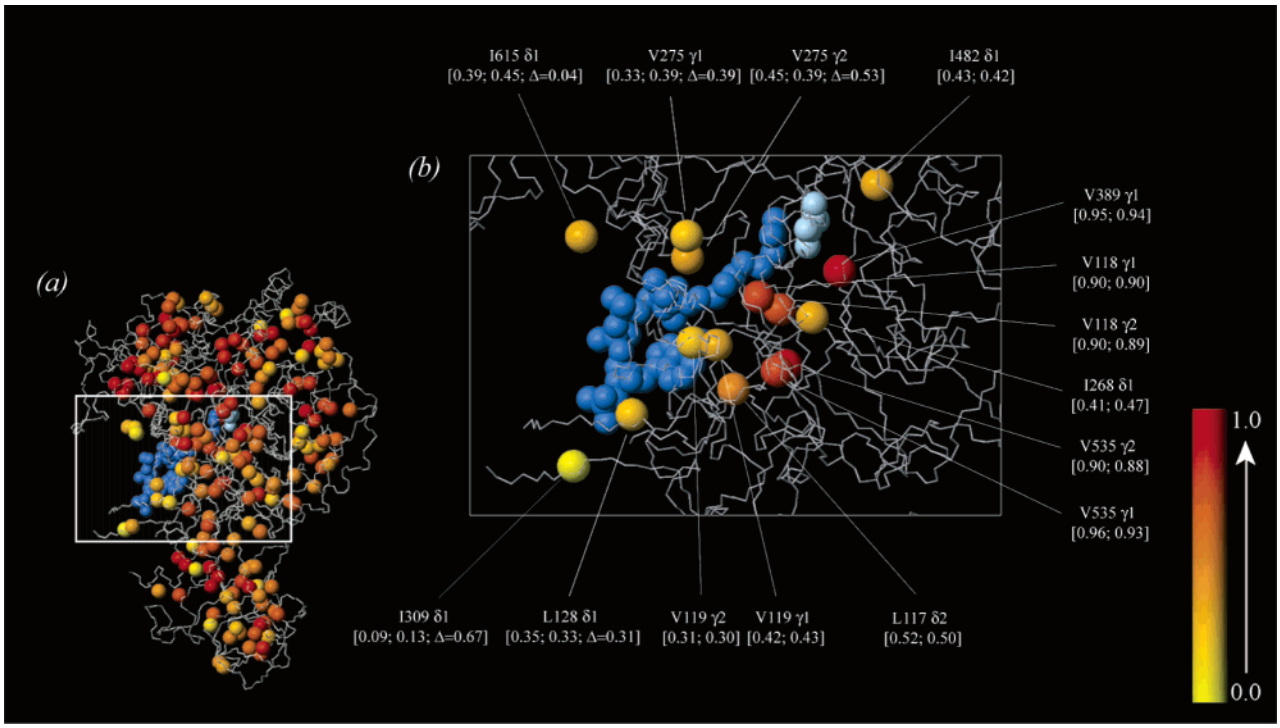

FIGURE 4: X-ray structure of MSG [PDB entry 1p7t (*15*)] with ILV methyls shown with spheres color-coded according to the values of ^{13}C-derived S^2_{axis} values for (a) the whole molecule and (b) the active site of MSG highlighting methyls whose ^{13}C atoms are located within 10 Å of one or more atoms of pyruvate (shown with blue-gray spheres) or acetyl-CoA (light blue spheres). The color coding (the scale shown at bottom right) was based on S^2_{axis} values obtained for the apo form of MSG, with ligands bound to MSG shown only to better delineate the binding site. The ^{13}C- and ^{2}H-derived S^2_{axis} values for the methyls in the proximity of the binding site (apo) in panel b are indicated in square brackets for each methyl site [S^2_{axis}(^{13}C),S^2_{axis}(^{2}H)]. Where values of S^2_{axis} ($\leq$1) are available for both apo and bound states of MSG, the difference is listed, Δ. This figure was prepared using MOLMOL (*49*).

^{13}C-derived S^2_{axis} values (apo form). The enzyme is comprised of four domains (*16*), including (i) a central β8/α8 (TIM barrel) core, with the eight β-strands forming a parallel β-sheet that wraps in a cylinder surrounded by eight α-helices, (ii) an N-terminal α-helical "clasp" linked to the TIM barrel by a long extended loop (residues 89−115), (iii) an α/β domain (residues 132−265, 296−333) appended to the core, and (iv) a C-terminal five-helix "plug" connected to the barrel by an extended loop (residues 550−588). It is clear that the methyls are distributed well over the entire protein, and in this respect, they serve as excellent probes of molecular dynamics. There is a quite uniform distribution in the level of dynamics observed throughout the protein. Nevertheless, several flexible areas do emerge, including the ligand binding interface, discussed below. In particular, regions spanning residues 152−155 (Val155) and 300−310 (Ile309 and Val310) in the α/β domain, residues in the linker between the core and the C-terminal plug (Val556 and Leu581), and residues in loops linking the four domains of the enzyme (Leu91, Val92, Ile260, and Ile265) exhibit significant mobility. Reduced ^{15}N $R_{1\rho}$ rates and ^{1}H−^{15}N NOE values of <0.6 have also been obtained for backbone amides in these regions (*18*).

Molecular Dynamics at the Binding Interface. As for all the enzymes sharing the TIM barrel fold, the active site of MSG is formed from residues at the C-termini of several β-strands of the core barrel and from the loop of the C-terminal domain spanning residues 614−631. The binding site of the acetyl-CoA substrate involves additionally the side chains of the β-hairpin loop of residues 300−310 and the first α-helix of the core barrel (*15*). Figure 4b highlights the region of the MSG molecule where acetyl-CoA (dark blue) and pyruvate (light blue) bind, with the ILV methyls whose ^{13}C atoms are within 10 Å of one or more atoms of the two bound ligands color-coded according to ^{13}C-derived S^2_{axis} values (apo form). Much, but not all, of the binding interface is surrounded by dynamic side chains that rigidify upon ligand binding (see the Supporting Information and Figure 4b). Included in the figure as well are the ^{13}C- and ^{2}H-derived S^2_{axis} values [$S^2_{axis}(^{13}C)$,$S^2_{axis}(^2H)$] for each of the proximal methyls (apo) and, where data are available for both free and bound states of the protein, the change in S^2_{axis} upon ligand binding, Δ. The level of agreement between order parameters obtained for ^{13}C and ^{2}H data is really quite remarkable considering the size of the system that is being studied.

A number of the methyl groups illustrated in Figure 4b have been shown to be dynamic on the millisecond time scale, including Leu117, Val119, Leu128, and Ile309 (*47*). Additionally, a comparison of ^{13}C $R_{1\rho}$ data recorded at 2.0 and 1.4 kHz spin lock field strengths also indicated possible micro- to millisecond time scale dynamics for Ile268 and for both methyls of Val275. Thus, for many residues at the binding interface, the dynamics span a wide range of time scales. Such micro- to millisecond contributions have little effect on measured ^{2}H spin relaxation rates since the deuterium quadrupolar interaction dominates the decay (compare the decay rates in panels c and d of Figure 1). However, conformational exchange processes can potentially lead to elevated ^{13}C $R_{1\rho}$ rates, resulting in increases in extracted $S^2_{axis}(^{13}C)$ values. The similarity between ^{2}H- and ^{13}C-derived order parameters for the residues in Figure 4b suggests that, at least for the methyls lining the binding site, exchange contributions are largely quenched by the 2 kHz ^{13}C spin lock field (see the discussion above).

A detailed account of all the interactions between MSG and its pyruvate and acetyl-CoA substrates has been presented by Anström et al. (*15*). Most of the interactions in the ternary pyruvate−acetyl-CoA−MSG complex are between polar side chains of MSG (guanidine of Arg, carboxyl of Asp, and hydroxyl groups of Ser and Tyr) and polar (amino and phosphate) groups of acetyl-CoA, and are in some cases mediated by several intervening water molecules. The changes in ^{13}C and ^{1}H chemical shifts and S^2_{axis} values upon formation of the ternary complex are summarized in the Supporting Information. Both ^{1}H and ^{13}C chemical shift changes qualitatively follow the pattern of ^{15}N shift changes reported earlier (*18*) and tend to concentrate around three main regions of the MSG structure involved in substrate binding, including (i) the loop connecting the first β1 strand and α1 helix of the core, (ii) the loop connecting strand β2 and helix α2 of the core as well as the β-hairpin of residues 300−311, and (iii) a pair of loops at the C-terminal end of the structure comprising residues 506−515 and 614−631. Changes in S^2_{axis} upon ligand binding are more prevalent throughout the sequence than might have been expected on the basis of chemical shift changes alone. On average, S^2_{axis} values increase in the ligated form of MSG by 0.018 (0.028) based on ^{13}C (^{2}H)-derived data. However, 52 (43) methyl sites exhibit a decrease in ^{13}C (^{2}H)-derived S^2_{axis} values in the ternary complex [$\Delta S^2_{axis}(^{13}C) = -0.054$; $\Delta S^2_{axis}(^2H) = -0.066$, on average]. Thus, the overall unfavorable entropic contributions to binding from decreased side chain flexibility are partially compensated by increases in entropy in regions of the molecule distinct from the binding site (see the Supporting Information). As expected, the most significant changes in methyl dynamics are observed close to the acetyl-CoA binding site. The especially noticeable "stiffening" of residues in the β-hairpin loop of residues 300−310 upon acetyl-CoA binding is consistent with results from our earlier ligand binding NMR study (*18*) and crystallographic data (*15*).

The changes in order parameters that accompany ligand binding can be recast in terms of changes in conformational entropy ΔS, using equations that relate the two (*50*−*52*). Assuming that the motion of each methyl 3-fold axis can be described in terms of a diffusion in a cone model, where every orientation is equiprobable (*51*), and that the nature of the motions does not change upon binding (only the cone angle varies, for example), changes in entropy between the apo and ligated forms of MSG can be calculated. Using a subset of 123 methyl groups, ΔS values of -180 ± 40 and -135 ± 30 J mol^{-1} K^{-1} are obtained from ^{2}H and ^{13}C measurements, respectively. To ensure that the same residues were included in the comparison, in the few cases ($\sim$10%) where values of $S^2_{axis}(^2H)$ were not available but where $S^2_{axis}(^{13}C)$ values were, we have used $S^2_{axis}(^{13}C)$ to calculate site-specific ΔS values and vice versa (3%). Notably, a significant portion of the entropy change, -52 ± 6 J mol^{-1} K^{-1} (^{2}H) and -53 ± 5 J mol^{-1} K^{-1} (^{13}C), derives from methyl sites that are proximal to the binding site (Figure 4b). In general, entropy changes calculated in this way must

be considered very approximate since they assume that the dynamics at each methyl site are uncorrelated, are based only on contributions from pico- to nanosecond time scale motions, and include contributions from only methyl groups for which measures are available in both apo and ligand-bound forms of the protein. Nevertheless, they indicate that changes in motion upon binding do contribute significantly (and unfavorably) to the association of MSG with its targets. In particular, the substantial contribution to the free energy of binding ($\Delta G = 16.4$ kJ/mol at 310 K) from changes in methyl dynamics at the binding interface suggests that taking the flexibility of this region into account might be very important in the design of potent inhibitors of this enzyme. For example, Figure 4b shows that one portion of the binding face is relatively rigid while the other regions are highly dynamic. A static structure that rules out certain potential ligands on the basis of steric clashes might be misleading in cases of dynamics, where, for example, regions of high plasticity could accommodate certain functionalities that, in the absence of motion, would be inaccessible (*1*). These functionalities might then provide significant enthalpic contributions to the free energy of binding that would stabilize the inhibitor−enzyme complex.

In summary, a ^{13}C and ^{2}H side chain dynamics study of MSG, an 82 kDa, 723-residue enzyme, has been presented. Excellent agreement between dynamics parameters obtained for both carbon- and deuterium-based approaches is obtained, establishing the robustness of the methodology. The MSG enzyme studied here is a member of a class of malate synthases that play an important role in a number of pathogenic microorganisms and for which inhibitors are sought. A mapping of the dynamics at the binding interface provides information that will be of value in achieving this goal.

SUPPORTING INFORMATION AVAILABLE

Description of experimental procedures employed for ^{2}H and ^{13}C relaxation measurements and theoretical considerations underlying data analyses. Included also are one figure showing the pulse schemes used for ^{13}C R_1 and $R_{1\rho}$ measurements, one figure showing the values of $S^2_{axis}(^{13}C)$ as a function of residue number and the changes in S^2_{axis} and methyl ^{1}H and ^{13}C chemical shifts occurring upon ternary complex formation, one table listing the ^{13}C- and ^{2}H-derived S^2_{axis} values in the apo and ligated states of MSG and the corresponding changes in site-specific entropy, and one table of the ^{13}C and ^{2}H spin relaxation rates. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Teague, S. J. (2003) Implications of protein flexibility for drug discovery, *Nat. Rev. Drug Discovery 2*, 527−541.
2. Homans, S. W. (2004) NMR spectroscopy tools for structure-aided drug design, *Angew. Chem., Int. Ed. 43*, 290−300.
3. Kay, L. E. (1998) Protein dynamics from NMR, *Nat. Struct. Biol. NMR Suppl. 5*, 513−516.
4. Ishima, R., and Torchia, D. A. (2000) Protein dynamics from NMR, *Nat. Struct. Biol. 7*, 740−743.
5. Palmer, A. G., Kroenke, C. D., and Loria, J. P. (2001) NMR methods for quantifying microsecond-to-millisecond motions in biological macromolecules, *Methods Enzymol. 339*, 204−238.
6. Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997) Attenuated T_2 relaxation by mutual cancellation of dipole−dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−12371.
7. Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E., and Kay, L. E. (2003) Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes, *J. Am. Chem. Soc. 125*, 10420−10428.
8. Miclet, E., Williams, D. C., Jr., Clore, G. M., Bryce, D. L., Boisbouvier, J., and Bax, A. (2004) Relaxation-optimized NMR spectroscopy of methylene groups in proteins and nucleic acids, *J. Am. Chem. Soc. 126*, 10560−10570.
9. Tugarinov, V., Ollerenshaw, J. E., and Kay, L. E. (2005) Probing side-chain dynamics in high molecular weight proteins by deuterium NMR spin relaxation: An application to an 82-kDa enzyme, *J. Am. Chem. Soc. 127*, 8214−8225.
10. Kronberg, H. L., and Krebs, H. A. (1957) Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle, *Nature 179*, 988−991.
11. Bishai, W. (2000) Lipid lunch for persistent pathogen, *Nature 406*, 683−685.
12. McKinney, J. D., Honer zu Bentrup, K., Munoz-Elias, E. J., Miczak, A., Chen, B., Chan, W. T., Swenson, D., Sacchettini, J. C., Jacobs, W. R. J., and Russell, D. G. (2000) Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase, *Nature 406*, 735−738.
13. Lorentz, M. C., and Fink, G. R. (2001) The glyoxalate cycle is required for fungal virulence, *Nature 412*, 83−86.
14. Factsheet No. 104 (2005) World Health Organization, http://www.who.int/mediacentre/factsheets/fs104/en/.
15. Anström, D. M., Kallio, K., and Remington, S. J. (2003) Structure of the *Escherichia coli* Malate Synthase G:pyruvate:acetyl-coenzyme A abortive ternary complex at 1.95 Å resolution, *Protein Sci. 12*, 1822−1832.
16. Howard, B. R., Endrizzi, J. A., and Remington, S. J. (2000) Crystal structure of *Escherichia coli* malate synthase G complexed with magnesium and glyoxylate at 2.0 Å resolution: Mechanistic implications, *Biochemistry 39*, 3156−3168.
17. Tugarinov, V., Choy, W. Y., Orekhov, V. Y., and Kay, L. E. (2005) Solution NMR-derived global fold of a monomeric 82-kDa enzyme, *Proc. Natl. Acad. Sci. U.S.A. 102*, 622−627.
18. Tugarinov, V., and Kay, L. E. (2003) Quantitative NMR studies of high molecular weight proteins: Application to domain orientation and ligand binding in the 723-residue enzyme malate synthase G, *J. Mol. Biol. 327*, 1121−1133.
19. Tugarinov, V., Muhandiram, R., Ayed, A., and Kay, L. E. (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: Chemical shift assignments and secondary structure of malate synthase G, *J. Am. Chem. Soc. 124*, 10025−10035.
20. Tugarinov, V., and Kay, L. E. (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods, *J. Am. Chem. Soc. 125*, 13868−13878.
21. Gardner, K. H., and Kay, L. E. (1997) Production and incorporation of ^{15}N, ^{13}C, ^{2}H (^{1}H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies, *J. Am. Chem. Soc. 119*, 7599−7600.
22. Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated ^{15}N-,^{13}C-,^{2}H-labeled proteins, *J. Biomol. NMR 13*, 369−374.
23. Tugarinov, V., and Kay, L. E. (2004) Stereospecific NMR assignments of prochiral methyls, rotameric states and dynamics of valine residues in malate synthase G, *J. Am. Chem. Soc. 126*, 9827−9836.
24. Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
25. Ishima, R., Louis, J. M., and Torchia, D. A. (1999) Transverse C-13 relaxation of CHD_2 methyl isotopomers to detect slow conformational changes of protein side-chains, *J. Am. Chem. Soc. 121*, 11589−11590.
26. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic relaxation in macromolecules: 1. Theory and range of validity, *J. Am. Chem. Soc. 104*, 4546−4559.
27. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic relaxation in macromolecules: 2. Analysis of experimental results, *J. Am. Chem. Soc. 104*, 4559−4570.

28. Kay, L. E., Bull, T. E., Nicholson, L. K., Griesinger, C., Schwalbe, H., Bax, A., and Torchia, D. A. (1992) The measurement of heteronuclear transverse relaxation times in AX_3 spin systems via polarization transfer techniques, *J. Magn. Reson. 100*, 538−558.
29. Abragam, A. (1961) *Principles of Nuclear Magnetism*, Clarendon Press, Oxford, U.K.
30. Mittermaier, A., and Kay, L. E. (1999) Measurement of methyl ^{2}H quadrupolar couplings in oriented proteins. How uniform is the quadrupolar coupling constant? *J. Am. Chem. Soc. 121*, 10608−10613.
31. Ishima, R., Petkova, A. P., Louis, J. M., and Torchia, D. A. (2001) Comparison of methyl rotation axis order parameters derived from model-free analyses of ^{2}H and ^{13}C longitudinal and transverse relaxation rates measured in the same protein sample, *J. Am. Chem. Soc. 123*, 6164−6171.
32. Tugarinov, V., Scheurer, C., Brüschweiler, R., and Kay, L. E. (2004) Estimates of methyl ^{13}C and ^{1}H CSA values ($\Delta\sigma$) in proteins from cross-correlated spin relaxation, *J. Biomol. NMR 30*, 397−406.
33. Richarz, R., Nagayama, K., and Wüthrich, K. (1980) Carbon-13 nuclear magnetic resonance relaxation studies of internal mobility of the polypeptide chain in basic pancreatic trypsin inhibitor and a selectively reduced analogue, *Biochemistry 19*, 5189−5196.
34. Henry, G. D., Weiner, J. H., and Sykes, B. D. (1986) Backbone dynamics of a model membrane protein: ^{13}C NMR spectroscopy of alanine methyl groups in detergent-solubilized M13 coat protein, *Biochemistry 25*, 590−598.
35. Nicholson, L. K., Kay, L. E., Baldisseri, D. M., Arango, J., Young, P. E., Bax, A., and Torchia, D. A. (1992) Dynamics of methyl groups in proteins as studied by proton-detected ^{13}C NMR spectroscopy. Application to the leucine residues of staphylococal nuclease, *Biochemistry 31*, 5253−5263.
36. Janin, J., Miller, S., and Chothia, C. (1988) Surface, subunit interfaces and interior of oligomeric proteins, *J. Mol. Biol. 204*, 155−164.
37. McCaldon, P., and Argos, P. (1988) Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences, *Proteins 4*, 99−122.
38. Tugarinov, V., and Kay, L. E. (2005) Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins, *Chembiochem 6*, 1567−1577.
39. Chaykovski, M. M., Bae, L. C., Cheng, M. C., Murray, J. H., Tortolani, K. E., Zhang, R., Seshadri, K., Findlay, J. H., Hsieh, S. Y., Kalverda, A. P., Homans, S. W., and Brown, J. M. (2003) Methyl side-chain dynamics in proteins using selective enrichment with a single isotopomer, *J. Am. Chem. Soc. 125*, 15767−15771.
40. Muhandiram, D. R., Yamazaki, T., Sykes, B. D., and Kay, L. E. (1995) Measurement of deuterium T_1 and $T_1\rho$ relaxation times in uniformly ^{13}C labeled and fractionally deuterium labeled proteins in solution, *J. Am. Chem. Soc. 117*, 11536−11544.
41. Kay, L. E., and Torchia, D. A. (1991) The effects of dipolar cross-correlation on ^{13}C methyl-carbon T_1, T_2 and NOE measurements in macromolecules, *J. Magn. Reson. 95*, 536−547.
42. Korzhnev, D. M., Skrynnikov, N. R., Millet, O., Torchia, D. A., and Kay, L. E. (2002) An NMR experiment for the accurate measurement of heteronuclear spin-lock relaxation rates, *J. Am. Chem. Soc. 124*, 10743−10753.
43. Ottiger, M., and Bax, A. (1999) How tetrahedral are methyl groups in proteins? A liquid crystal NMR study, *J. Am. Chem. Soc. 121*, 4690−4695.
44. Mittermaier, A., and Kay, L. E. (2002) Effect of deuteration on some structural parameters of methyl groups in proteins as evaluated by residual dipolar couplings, *J. Biomol. NMR 23*, 35−45.
45. Best, R. B., Clarke, J., and Karplus, M. (2004) The origin of protein side-chain order parameter distributions, *J. Am. Chem. Soc. 126*, 7734−7735.
46. Mittermaier, A., Kay, L. E., and Forman-Kay, J. D. (1999) Analysis of deuterium relaxation-derived methyl axis order parameters and correlation with local structure, *J. Biomol. NMR 13*, 181−185.
47. Korzhnev, D. M., Kloiber, K., Kanelis, V., Tugarinov, V., and Kay, L. E. (2004) Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: Application to a 723-residue enzyme, *J. Am. Chem. Soc. 126*, 3964−3973.
48. Beckmann, P. A., Buser, C. A., Mallory, C. W., Mallory, F. B., and Mosher, J. (1998) Methyl reorientation in solid 3-ethylcrysene and 3-isopropylcrysene, *Solid State Nucl. Magn. Reson. 12*, 251−256.
49. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
50. Akke, M., Brüschweiler, R., and Palmer, A. (1993) NMR order parameters and free energy: An analytic approach and application to cooperative calcium binding by calbindin D9k, *J. Am. Chem. Soc. 115*, 9832−9833.
51. Yang, D., and Kay, L. E. (1996) Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: Application to protein folding, *J. Mol. Biol. 263*, 369−382.
52. Lee, A. L., Kinnear, S. A., and Wand, A. J. (2000) Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex, *Nat. Struct. Biol. 7*, 72−77.

BI0519809