J|A|C|S COMMUNICATIONS

Published on Web 05/13/2010

Determination of Isoleucine Side-Chain Conformations in Ground and Excited States of Proteins from Chemical Shifts

D. Flemming Hansen,* Philipp Neudecker, and Lewis E. Kay*

University of Toronto, Departments of Molecular Genetics, Biochemistry, and Chemistry, 1 King's College Circle, Toronto, Ontario, Canada

Received March 11, 2010; E-mail: kay@pound.med.utoronto.ca; flemming@pound.med.utoronto.ca

Over the past decade it has become increasingly clear that protein backbone chemical shifts are strongly linked to molecular conformation[1] and that, importantly, they can be used as the sole experimental input to obtain accurate three-dimensional structures of small rigid proteins using database driven computational approaches.[2,3] In addition, it has been established that such chemical shifts provide a qualitative measure of dynamics along the protein backbone.[4,5] Thus, significant effort is now focused on establishing quantitative structure/dynamics−chemical shift relationships for side-chain nuclei, similar to those that have been shown to be so valuable in the context of the protein backbone.[6] One example in this regard is the recent demonstration that the chemical shifts of Leu side-chain methyl carbons correlate with conformational averaging of the χ_2 side-chain dihedral angle and that such shifts can be used to define the (χ_1,χ_2) rotameric states for this residue.[7,8] The utility of chemical shifts in general is particularly noteworthy given that they are the NMR parameter that is both the most easily accessed and the most accurately measured. Their importance is further established in studies of the structures and dynamics of low populated, transiently formed excited states, as we will show below, since the availability of other experimental probes of structure and motion in these systems is often limited.

Below we build upon earlier work focusing on Leu side-chain[7,9] by deriving a very simple relationship between the Ile $C^{\delta1}$ chemical shift and the χ_2 dihedral angle that can be used to provide a direct measure of conformational sampling of Ile side-chains from measured $C^{\delta1}$ shifts. Because methyl chemical shifts can be measured with high sensitivity, it becomes possible to obtain such dynamics information on large macromolecular machines and on molecular complexes that may not be amenable to study using other NMR approaches, such as those involving measurement of scalar couplings. Furthermore, because $C^{\delta1}$ chemical shifts can be measured in invisible, low populated states of proteins by relaxation dispersion (RD) NMR experiments so long as (i) they are populated to greater than ∼0.5% and (ii) they interconvert with a visible ground state on the millisecond time scale, quantifying Ile side-chain structure and dynamics in excited states becomes possible.

The Ile side-chain χ_2 dihedral angle can sample four distinct conformations (trans, gauche−, gauche+, gauche100) with populations of (∼81%, ∼15%, ∼2%, ∼2%),[10] based on analysis of a set of high resolution crystal structures and using methyl groups with *B*-factors less than 20 Å^2, Figure 1. Within each of these four conformations the range of χ_2 angles varies by no more than ∼8°, strongly suggesting that motions about the χ_2 angle are well-described by jumps between rotameric states and that conformational sampling of χ_2 can be described in terms of the populations of the individual rotamers. To very good approximation only the trans and gauche− conformers are sampled in solution (see also Supporting Information, SI, establishing that the Gauche100 conformation is not populated), and in what follows only these rotameric states are considered.

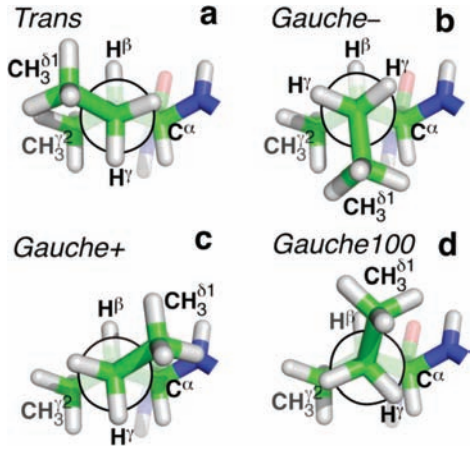

Figure 1. Four Ile conformations present in a database of high-resolution crystal structures.[10] The average χ_2 dihedral angles are (a) Trans: $\chi_2 = 170°$ (b) Gauche−: $\chi_2 = 300°$, (c) Gauche+: $\chi_2 = 66°$ (d) Gauche100: $\chi_2 = 100°$.

The chemical shift of an aliphatic carbon depends on neighboring substituents and on steric effects.[11] Of interest here is the so-called γ-gauche effect whereby the chemical shift of a given ^{13}C nucleus is significantly influenced by its position relative to γ-substituents (see below). Such an affect was recently exploited in the determination of Leu side-chain (χ_1,χ_2) rotamer populations in ground and excited states of proteins from differences in $^{13}C^{\delta1}$, $^{13}C^{\delta2}$ chemical shifts.[7,9] Following empirical rules for the γ-gauche effect (e.g., ref 11) one would predict that, in the case of Ile, $\delta(C^{\delta1})_{trans} \approx \delta(C^{\delta1})_{gauche+}$, where $\delta(C^{\delta1})_x$ is the chemical shift of $^{13}C^{\delta1}$ in conformation *x*, since a single γ-gauche effect (involving $^{13}C^{\gamma2}$ for trans, for example) is operative. In contrast, two such effects, from both $C^{\gamma2}$ and C^{α}, are present for gauche−, resulting in a decreased shift. These predictions were verified through a series of DFT calculations on a small model complex (see SI) showing further that $\delta(C^{\delta1})_{trans} - \delta(C^{\delta1})_{gauche-} \approx 5.9$ ppm, very similar to what is observed experimentally (see below). An important consequence is that the population of the gauche− conformation can be calculated directly from the chemical shift of $C^{\delta1}$:

$$\delta(C^{\delta1}) \approx \delta(C^{\delta1})_{gauche-}p_{gauche-} + \delta(C^{\delta1})_{trans}(1 - p_{gauche-}) \quad (1)$$

The γ-gauche effect is so large that it can be measured easily from simple 2D $^1H-^{13}C$ correlation spectra; small errors in referencing as well as ring current shifts will not obscure the effect. A series of calculations using a set of six proteins (100 Ile residues;

see SI) establishes that the contributions that ring currents make to measured $^{13}C^{\delta1}$ shifts in proteins, δ_{RC}, are in general rather small, with $\langle|\delta_{RC}|\rangle = 0.13$ ppm $\pm$ 0.17 ppm and a smallest(largest) calculated value of $-0.94(+0.43)$ ppm.

Prior to using eq 1 to estimate values of $p_{gauche-}$, $p_{trans} = (1 - p_{gauche-})$, it is necessary to first obtain estimates for $\delta(C^{\delta1})_{gauche-}$ and $\delta(C^{\delta1})_{trans}$. Here we have used a database of seven proteins (see SI) for which both $\delta(C^{\delta1})$ chemical shifts and $^3J(C^{\delta1},C^{\alpha})$ scalar couplings have been measured (49 Ile). It is well-known that three-bond scalar couplings are sensitive to the intervening dihedral angle so that Ile χ_2 can be probed by measurement of the $^3J(C^{\delta1},C^{\alpha})$ scalar coupling via

$$^3J(C^{\delta1},C^{\alpha}) = {}^3J_{trans}(C^{\delta1},C^{\alpha})p_{trans} + {}^3J_{gauche-}(C^{\delta1},C^{\alpha})(1 - p_{trans}) \quad (2)$$

where $^3J_{trans}(C^{\delta1},C^{\alpha}) \approx 3.7$ Hz and $^3J_{gauche-}(C^{\delta1},C^{\alpha}) \approx 1.5$ Hz.[12] Equations 1 and 2 predict a linear correlation between $\delta(C^{\delta1})$ and $^3J(C^{\delta1},C^{\alpha})$ (see SI), so that values for $\delta(C^{\delta1})_{gauche-}$ and $\delta(C^{\delta1})_{trans}$ can be obtained from the extreme ends of the correlation plot. Figure 2a shows the derived correlation. As expected the majority of the Ile χ_2 angles in the seven proteins are close to trans (T) and the positions of 47 of the 49 experimental data points (red) are well described by averaging between the gauche$-$ (G$-$) and T conformations. The two outliers (black) derive from residues with χ_2 in the rare gauche+ (G+) conformation,[10] since large $\delta(C^{\delta1})$ and small $^3J(C^{\delta1},C^{\alpha})$ values are predicted for this rotamer (see SI). A G+ conformation was further verified for I104 from maltose binding protein[13] that contributes one of the G+ outliers (asterisk) on the basis of the large $^3J(C^{\delta1},C^{\gamma2})$ scalar coupling ≈ 3.2 Hz.

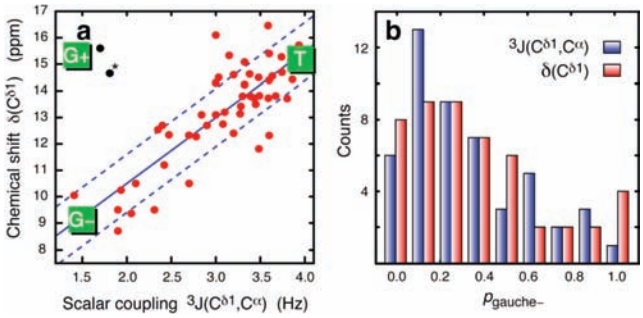

Figure 2. (a) Correlation between the three-bond $C^{\delta1}-C^{\alpha}$ scalar coupling, $^3J(C^{\delta1},C^{\alpha})$, and the $C^{\delta1}$ chemical shift, $\delta(C^{\delta1})$. The blue line is the best-fit line to the red points, $\delta(C^{\delta1}) = 2.5 \times {}^3J(C^{\delta1},C^{\alpha})$ ppm/Hz + 5.5 ppm, with an rmsd of 1.1 ppm. The dashed blue lines are the best-fit line $\pm$ rmsd. (b) Histogram of the derived populations of the gauche$-$ conformation using the data of (a), omitting the two G+ points (black), with $p_{gauche-}$ calculated from eq 3 (red) or eq 2 (blue).

The correlation between $^3J(C^{\delta1},C^{\alpha})$ and $\delta(C^{\delta1})$ in Figure 2a establishes that $\delta(C^{\delta1})$ provides a relatively good measure of the Ile side-chain χ_2 conformation. Assuming $^3J_{trans}(C^{\delta1},C^{\alpha}) = 3.7$ Hz and $^3J_{gauche}(C^{\delta1},C^{\alpha}) = 1.5$ Hz, Figure 2a predicts values of $\delta(C^{\delta1})_{gauche-} = 9.3$ ppm and $\delta(C^{\delta1})_{trans} = 14.8$ ppm, in good agreement with 8.8 ppm (G$-$) and 14.7 ppm (T) obtained from DFT (see SI). Using the experimentally determined $\delta(C^{\delta1})_{gauche-}$ and $\delta(C^{\delta1})_{trans}$ values the following relationship between $p_{gauche-}$ and $\delta(C^{\delta1})$ is obtained

$$p_{gauche-} \approx \begin{cases} 1 \text{ for } \delta(C^{\delta1}) < 9.3\text{ppm} \\ 0 \text{ for } \delta(C^{\delta1}) > 14.8\text{ppm} \\ \text{else} \\ [14.8\text{ppm} - \delta(C^{\delta1})]/5.5\text{ppm} \end{cases} \quad (3)$$

The rmsd between the red points of Figure 2a and the best-fit line (solid blue) (1.1 ppm) translates into a maximum uncertainty of approximately 0.2 in $p_{gauche-}$ derived from eq 3; however this assumes no error in experimental $^3J(C^{\delta1},C^{\alpha})$ values and that they translate into exact populations of rotameric states via eq 2. The actual uncertainty in $p_{gauche-}$ values derived from eq 3 is therefore expected to be significantly lower than 0.2. It is worth noting that the rmsd value of 1.1 ppm in Figure 2a is of the same order as the accuracy of predicted α- and β-carbon chemical shifts from three-dimensional structures.[6,14] Prediction of $C^{\delta1}$ chemical shifts, however, is predicated only on the χ_2 angle.

Figure 2b shows histograms of $p_{gauche-}$ values derived from either $^3J(C^{\delta1},C^{\alpha})$ (blue) or $\delta(C^{\delta1})$ (red) for the 47 (red) points in Figure 2a. Averaged over all values, $\langle p_{gauche-}\rangle_{^3J} = 29\%$, $\langle p_{gauche-}\rangle_{\delta} = 32\%$, which agrees well with $p_{gauche-} = 35\%$ calculated from the Ile $\delta(C^{\delta1})$ random coil value of 12.9 ppm[15] (eq 3). Although consistent $\langle p_{gauche-}\rangle$ values are obtained from both scalar couplings and chemical shifts that closely predict Ile $C^{\delta1}$ random coil shifts measured for small peptides[15] and unfolded proteins,[16] $\langle p_{gauche-}\rangle$ values from NMR and X-ray do differ somewhat ($\sim$30% vs 15%). This may reflect subtle differences between conformations in the solution and solid state as well as the significant differences in the temperatures at which the two classes of experiment are performed.

Equation 3 provides a relation for estimating the extent of conformational sampling of Ile χ_2 purely from measurement of $C^{\delta1}$ chemical shift values under the assumption that only T and G$-$ rotamers are populated. In cases where $^3J(C^{\delta1},C^{\alpha})$ values can be measured as well it, is possible to identify the small fraction (2%) of Ile residues for which χ_2 is in the G+ conformation. Lack of coupling data is not a deterrent, however, since in over 95% of the cases the conformation is a weighted average between the T and G$-$ rotameric states.

As an illustration of the method consider the A39V/N53P/V55L mutant SH3 domain from the Fyn tyrosine kinase, which folds on the millisecond time scale via an on-pathway intermediate, U$\leftrightarrow$I$\leftrightarrow$F. The folding of this domain has been quantified by RD NMR spectroscopy, and at temperatures below 25 °C dispersion profiles can be interpreted as arising from a global exchange event involving the interconversion between F and I states with negligible accumulation of U ($\sim$2% I).[17] As in other cases where invisible states are probed by RD, chemical shifts and residual dipolar couplings are to this point the only primary experimental data that are accessible.[18] Figure 3 shows RD profiles of the two Ile residues in the SH3 domain (28 and 50) recorded at static magnetic field strengths corresponding to ^{1}H resonance frequencies of 500 MHz

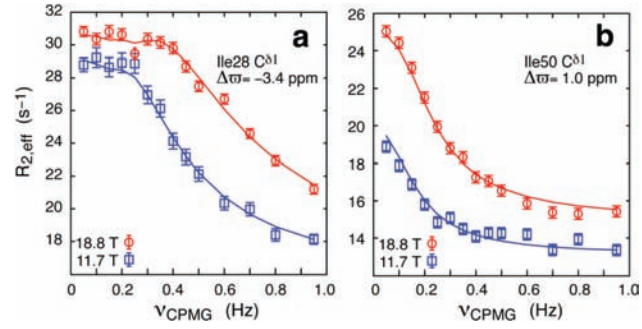

Figure 3. Methyl CPMG relaxation dispersion profiles for I28 $C^{\delta1}$ (a) and I50 $C^{\delta1}$ (b) of the A39V/N53P/V55L mutant Fyn SH3 domain recorded at two magnetic field strengths, 18.8 T (red) and 11.7 T (blue), 20 °C. The population of the intermediate state, $p_I = 1.9 \pm 0.1\%$, and the exchange rate for the interconversion of F and I, $k_{ex,FI} = 775 \pm 15$ s^{-1}, along with values of $|\Delta\varpi|$ were obtained from fits of the profiles as described previously.[17,20] Experimental details are given in the Supporting Information.

Table 1. Chemical Shifts and Gauche− Populations of the Folded and Intermediate States of A39V/N53P/V55L Fyn SH3, 20 °C

	I28 $C^{\delta 1}$		I50 $C^{\delta 1}$	
	Folded	Intermediate	Folded	Intermediate
$\delta(C^{\delta 1})$	10.3 ppm	13.7 ppm	14.8 ppm	13.8 ppm
$p_{\text{gauche}-}$	0.82	0.20	∼0	0.18

(blue) and 800 MHz (red). Fits of these curves (along with those from other methyl groups) to a two-site model of exchange (between I and F) provides residue specific $|\Delta\varpi|$ values, where $\Delta\varpi$ is the difference between chemical shifts in states I and F. Signs of $\Delta\varpi$ values have been obtained from ^{1}H,^{13}C HSQC/HMQC data sets recorded at different magnetic field strengths[19] so that $\delta^{I}(C^{\delta 1}) = \delta^{F}(C^{\delta 1}) - \Delta\varpi$ can be calculated, Table 1, where $\delta^{I}(C^{\delta 1})$ and $\delta^{F}(C^{\delta 1})$ are the $^{13}C^{\delta 1}$ chemical shifts in states I and F, respectively. Once $\delta^{I}(C^{\delta 1})$ values are known $p_{\text{gauche}-}$ values in the I state are readily obtained from eq 3. Note that eq 3 is equally valid for well structured, partially structured, or disordered proteins so long as the dynamics about the χ_2 rotamer can be correctly described by the interconversion between predominant gauche− and trans conformers, as described above.

I28 and I50 of the Fyn SH3 domain are located in strands $\beta 2$ and $\beta 4$, respectively.[17] In the folded state I28 packs against the side-chain of F4 in the small β-sheet formed by strands $\beta 1$ and $\beta 5$, so that $\Delta\varpi$ for this residue can be used as a qualitative indicator of the presence of native-like interactions between $\beta 2$ and the $\beta 1$:$\beta 5$ sheet. By contrast, I50 points into the core of the protein in the folded state so that $\Delta\varpi$, $p_{\text{gauche}-,I}$ values for this residue serve as a qualitative probe of core packing in the intermediate. A quantitative analysis of the RD data establishes that χ_2 of I28 changes significantly between states F and I, from 82% gauche− to 20% gauche−. In contrast, the I50 χ_2 changes much less, from ∼0% gauche− (F) to 18% gauche− (I), indicating that upon interconversion from F to I the I50 χ_2 rotamer conformer shifts from essentially pure trans to one that involves some degree of averaging, although still predominantly trans. It is noteworthy that the χ_2 rotameric states for the F state obtained from eq 3 are in excellent agreement with the crystal structure.[17]

Overall, the significant conformational change of I28 is consistent with a non-native interaction between $\beta 2$ and the $\beta 1$:$\beta 5$ sheet in state I, while the I50 probe indicates that the core of the folding intermediate is only slightly more dynamic than the folded state, consistent with previous results showing that much of the core is already formed in the intermediate state.[17] These results are in agreement with a previous RD analysis of Leu side-chain conformations in the A39V/N53P/V55L Fyn SH3 domain showing that in many cases rotameric states in the intermediate are similar to those in the folded state, with some conformations shifted slightly toward random coil, except for L55, located in strand $\beta 5$, whose conformation is random coil in the intermediate state.[9]

In summary, it has been shown that conformational sampling of the Ile side-chain χ_2 angle can be obtained directly from measurement of the Ile $^{13}C^{\delta 1}$ chemical shift. Chemical shift measurements are very sensitive and particularly so for methyl groups where accurate shifts can be obtained in complexes with molecular weights in the 1 MDa range[21] so long as appropriate labeling and pulse sequence methodology are employed. It is also possible to measure very accurate ^{13}C methyl chemical shifts in invisible excited states by RD NMR spectroscopy, as demonstrated here and in other applications.[22] These shifts are sensitive to motions in the invisible state that occur on a time scale which is fast compared to the millisecond interconversion process that is responsible for the dispersion profiles in the first place. In the present study Ile χ_2 rotamer populations have been determined that report on Ile side-chain structure and dynamics and that can be directly used to obtain changes in Ile side-chain conformational entropy between exchanging states. The entropy derived from chemical shifts reports on motions covering a broad spectrum of time scales, ranging from picoseconds to near milliseconds, that frequently encompasses the time frame for many important biological processes, including enzyme reactions, ligand binding events, and protein folding. The methodology presented thus extends the utility of chemical shifts as probes of structure and dynamics to include Ile side-chain and in the context of RD studies provides an avenue for characterizing Ile side-chain dynamics in invisible excited states at a level of detail that has been, until now, reserved for applications to visible, ground states of proteins.

Acknowledgment. This work was supported by a grant from the Canadian Institutes of Health Research (CIHR) to L.E.K. D.F.H. acknowledges CIHR for a postdoctoral fellowship. L.E.K. holds a Canada Research Chair in Biochemistry.

Supporting Information Available: Details of DFT calculations, data used to produce Figure 2, protein production and NMR data analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289–302.
(2) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615–9620.
(3) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685–4690.
(4) Berjanskii, M. V.; Wishart, D. S. *J. Biomol. NMR* **2008**, *40*, 31–48.
(5) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. *J. Biomol. NMR* **2009**, *44*, 213–23.
(6) Shen, Y.; Bax, A. *J. Biomol. NMR* **2007**, *38*, 289–302.
(7) Mulder, F. A. A. *ChemBioChem* **2009**, *10*, 1477–9.
(8) London, R. E.; Wingad, B. D.; Mueller, G. A. *J. Am. Chem. Soc.* **2008**, *130*, 11097–105.
(9) Hansen, D. F.; Neudecker, P.; Vallurupalli, P.; Mulder, F. A. A.; Kay, L. E. *J. Am. Chem. Soc.* **2010**, *132*, 42–3.
(10) Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C. *Proteins* **2000**, *40*, 389–408.
(11) Williams, D. H.; Fleming, I. *Spectroscopic methods in organic chemistry*, 5th ed.; The McGraw-Hill Companies: London, 1995.
(12) Chou, J. J.; Case, D. A.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 8959–66.
(13) Gardner, K. H.; Zhang, X.; Gehring, K.; Kay, L. E. *J. Am. Chem. Soc.* **1998**, *120*, 11738–11748.
(14) Neal, S.; Nip, A. M.; Zhang, H.; Wishart, D. S. *J. Biomol. NMR* **2003**, *26*, 215–40.
(15) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67–81.
(16) Korzhnev, D. M.; Mittermaier, A. K.; Kay, L. E. *J. Biomol. NMR* **2005**, *31*, 337–42.
(17) Neudecker, P.; Zarrine-Afsar, A.; Choy, W.-Y.; Muhandiram, D. R.; Davidson, A. R.; Kay, L. E. *J. Mol. Biol.* **2006**, *363*, 958–76.
(18) Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J. Biomol. NMR* **2008**, *41*, 113–20.
(19) Skrynnikov, N. R.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 12352–60.
(20) Hansen, D. F.; Vallurupalli, P.; Lundstrom, P.; Neudecker, P.; Kay, L. E. *J. Am. Chem. Soc.* **2008**, *130*, 2667–75.
(21) Sprangers, R.; Kay, L. E. *Nature* **2007**, *445*, 618–22.
(22) Lundstrom, P.; Vallurupalli, P.; Religa, T. L.; Dahlquist, F. W.; Kay, L. E. *J. Biomol. NMR* **2007**, *38*, 79–88.

JA102090Z