Supporting Material

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.

Ring Current Effects:

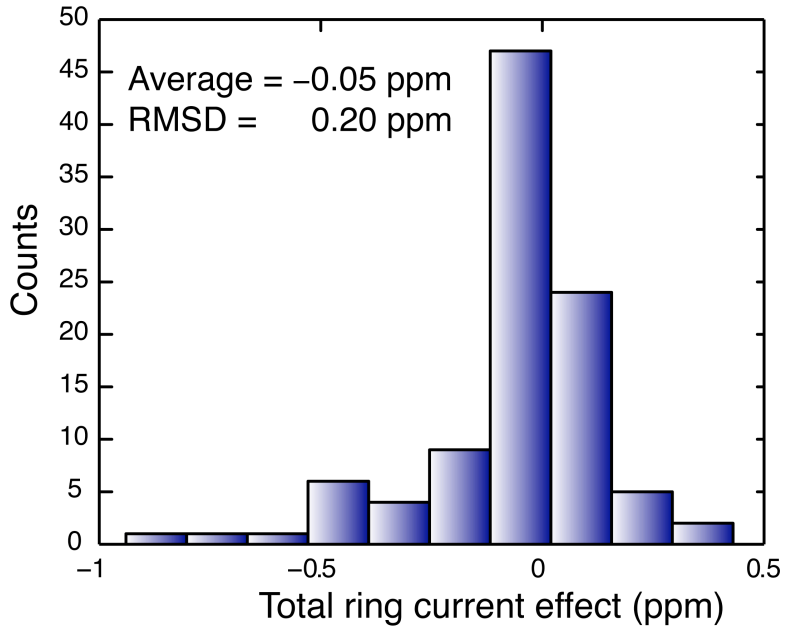

Figure S1: Histogram of ring current effects calculated for $C^{\delta 1}$ nuclei of T4 lysozyme L99A (pdb: 3dmv), maltose binding protein (pdb: 3hpi), malate synthase G (pdb: 2jqx), C-terminal SH2 domain of phospholipase C-γ1 (pdb: 2pld), B1 immunoglobulin binding domain of peptostreptococcal protein L (pdb: 1hz6) and human ubiquitin (pdb: 1ubq). The ring current effects were calculated using a modified version of ShiftX[1] that includes the $C^{\delta 1}$ nucleus of Ile residues (program available upon request).

Density Functional Theory Calculations of $^{13}C^{\delta 1}$ Chemical Shifts:

All density functional theory calculations described below were carried out with Gaussian 03 Rev. B.05[2].

Generation of model complexes

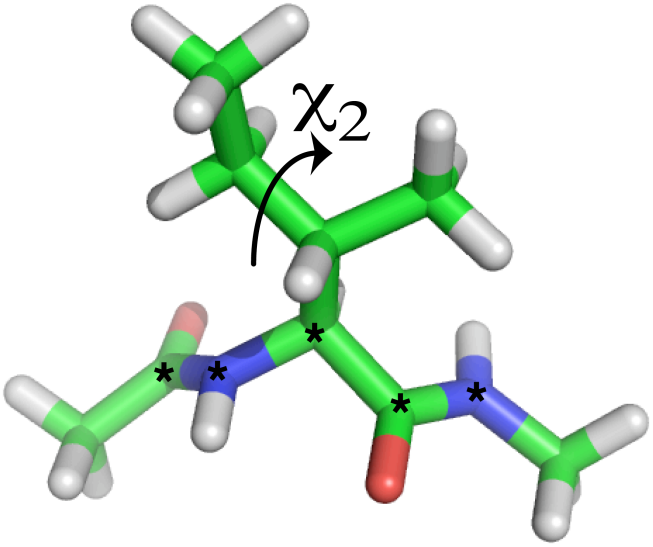

Figure S2. Structure of the model complex used for calculating chemical shifts of the $C^{\delta 1}$ atom as a function of χ_2 (Figure S3). The model complex was generated from the N59-I60-K61 fragment of the protein L crystal structure (pdb: 1hz6) and modified with Molden 4.8 (http://www.cmbi.ru.nl/molden/) to generate the final complex.

Initial geometry optimizations

An initial geometry optimization of the model complex structure was performed using the three-parameter Becke-Lee-Yang-Parr (B3LYP) exchange-correlation-functional in the 6-31G*[3,4] basis set (resulting in 231 basis functions). Initially, $(\chi_1,\chi_2)=(300^o,170^o)$ and the backbone was placed in a β-sheet extended conformation. The χ_2 dihedral angle and all atoms marked with * in Figure S2 were frozen during the geometry optimization. The structure shown in Figure S2 is the result of the initial optimization, where the χ_1 angle has changed slightly to 301.1^o.

Model complexes with different χ_2 dihedral angles χ_2={10, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 110, 130, 150, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 230, 250, 270, 280, 285, 290, 295, 300, 305, 310, 315, 320, 330, 350} degrees were generated from the initially optimized model using an in-house written program and structures were subsequently optimized using a B3LYP functional in the 6-31G* basis set, fixing the χ_2 dihedral angle and the atoms marked with * in Figure S2.

Calculation of chemical shifts

Chemical shifts were calculated using the geometry optimized structures of each of the 39 model complexes using the EPR-III[5] basis on all atoms (resulting in 718 basis functions). The EPR-III basis is particularly advantageous for calculating chemical shifts since it has been optimized to enhance the description of electron density in the nuclear region. Chemical shifts were calculated using the gauge-independent atomic orbital (GIAO)[6] approach implemented in Gaussian 03 (see Figure S3).

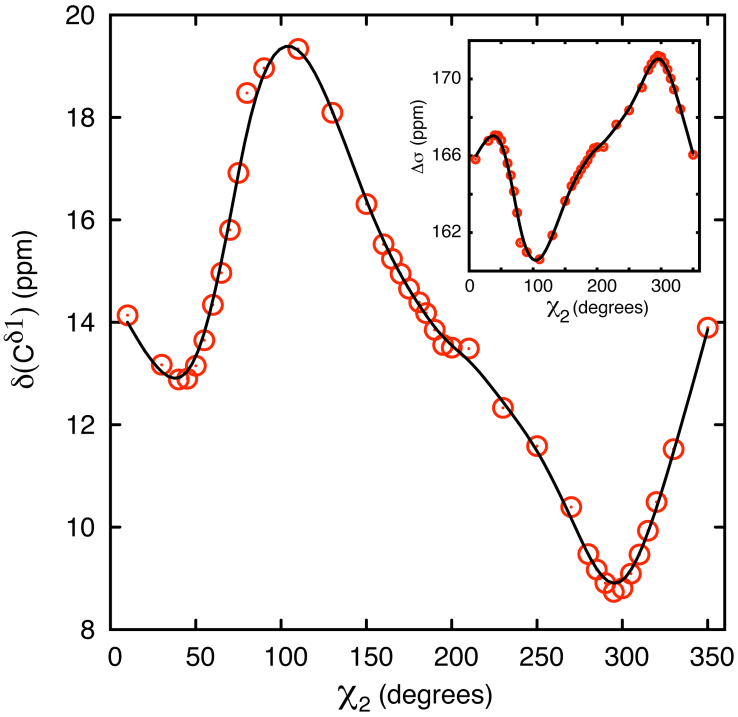

Figure S3. Calculated chemical shifts, $\delta(C^{\delta 1})$, and shieldings, $\Delta\sigma$, for the 39 model complexes with different χ_2 angles, geometrically optimized as described above. The continuous black line is a smooth interpolation between the individual calculated points that are shown in red. The chemical shifts $\delta(C^{\delta 1})$ were calculated as $\delta = -(\Delta\sigma - \Delta\sigma_{ref})$, where $\Delta\sigma_{ref}$ was determined so that the chemical shift corresponding to random coil $\delta_{RC} = 0.32\delta_{300} + 0.04\delta_{60} + 0.64\delta_{170}$ agrees with the experimentally observed random coil chemical shift of 12.9ppm[7], where δ_x is the chemical shift calculated for $\chi_2 = x$. The coefficient 0.32 was obtained from $\langle p_{gauche-} \rangle_\delta = 32\%$ (Figure 2b), while the coefficient 0.04 derives from the fact that 2 of the 49 residues used in our database were in the G+ conformation (Figure 2a).

Derivation of Eq 3

By assuming that trans and gauche– are the only populated rotameric states the population $p_{gauche-}$ can be isolated from eq 2:

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

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

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

$$p_{gauche-} = \frac{{}^3J(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})}{{}^3J_{gauche-}(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})} \qquad \text{(S1)}$$

Inserting Eq. S1 into Eq 1 gives:

$$\begin{aligned}
\delta(C^{\delta1}) &= (\delta(C^{\delta1})_{gauche-} - \delta(C^{\delta1})_{trans})p_{gauche-} + \delta(C^{\delta1})_{trans} \\
&= \frac{{}^3J(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})}{{}^3J_{gauche-}(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})}(\delta(C^{\delta1})_{gauche-} - \delta(C^{\delta1})_{trans}) + \delta(C^{\delta1})_{trans} \\
&= {}^3J(C^{\delta1},C^{\alpha})\frac{\delta(C^{\delta1})_{gauche-} - \delta(C^{\delta1})_{trans}}{{}^3J_{gauche-}(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})} \\
&\quad - {}^3J_{trans}(C^{\delta1},C^{\alpha})\frac{\delta(C^{\delta1})_{gauche-} - \delta(C^{\delta1})_{trans}}{{}^3J_{gauche-}(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})} + \delta(C^{\delta1})_{trans} \quad ,
\end{aligned}$$

Thus, the correlation in Figure 2a is given by

$$\text{slope} = \frac{\delta(C^{\delta1})_{gauche-} - \delta(C^{\delta1})_{trans}}{{}^3J_{gauche-}(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})}$$

$$\text{intercept} = -{}^3J_{trans}(C^{\delta1},C^{\alpha})\frac{\delta(C^{\delta1})_{gauche-} - \delta(C^{\delta1})_{trans}}{{}^3J_{gauche-}(C^{\delta1},C^{\alpha}) - {}^3J_{trans}(C^{\delta1},C^{\alpha})} + \delta(C^{\delta1})_{trans}$$

Using ${}^3J_{trans}(C^{\delta1},C^{\alpha}) \approx 3.7$ Hz and ${}^3J_{gauche}(C^{\delta1},C^{\alpha}) \approx 1.5$ Hz[8] one obtains that

$$\delta(C^{\delta1})_{trans} - \delta(C^{\delta1})_{gauche-} = 5.5 \text{ ppm}$$

$$\delta(C^{\delta1})_{trans} = 14.8 \text{ ppm}$$

The Gauche100 Conformation is Not Populated in Solution

From the DFT calculations shown above we found that $\delta(C^{\delta1})_{gauche100} - \delta(C^{\delta1})_{trans} \approx 4.5$ ppm. In Figure 2 we observe no experimental data point with a chemical shift of ~ 19 ppm and a small coupling constant, which strongly suggests that the gauche100 conformation that is observed in 2% of the crystal structures is not present in the solution state.

Protein Sample Conditions

A sample of the A39V/N53P/V55L mutant of the Fyn SH3 domain was prepared by protein over-expression in M9 minimal media with 3g/L [1-^{13}C]-glucose as the sole carbon source, essentially as described in detail previously[9,10]. The sample was 1.0 mM in protein, 0.2 mM EDTA, 0.05% NaN_3, 50 mM sodium phosphate, 90%/10% H_2O/D_2O, pH 7.0. Also, a sample of [^{1}H,^{15}N,^{13}C]-L99A T4 lysozyme was prepared as described previously[11]. The sample was 1.1 mM in protein, 50 mM sodium phosphate, 25 mM NaCl, 2 mM EDTA, and 2 mM NaN_3 at pH 5.5.

NMR Spectroscopy and Data Analysis

Relaxation dispersion measurements on A39V/N53P/V55L Fyn SH3

Single quantum ^{13}C relaxation dispersion profiles[12] were recorded at 20 ºC and at magnetic fields strengths of 11.4 and 18.8 T for the A39V/N53P/V55L Fyn SH3 domain. Relaxation dispersion profiles, $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}})$, were generated from peak intensities, $I_1(\nu_{\mathrm{CPMG}})$, in a series of 2D ^{1}H-^{13}C correlation maps measured as a function of CPMG frequency, $\nu_{\mathrm{CPMG}} = 1/(4\tau_{\mathrm{CP}})$, where $2\tau_{\mathrm{CP}}$ is the interval between consecutive 180º-refocusing pulses of the CPMG sequence. Effective relaxation rates were calculated as $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}}) = \ln(I_0/I_1(\nu_{\mathrm{CPMG}}))/T_{\mathrm{relax}}$, where I_0 is the peak intensity in a reference spectrum recorded without the relaxation delay T_{relax}[13]. Dispersion profiles were recorded with T_{relax} = 40 ms and 14 values of ν_{CPMG} ranging from 50-950 Hz, along with 2 repeats for error analysis.

All data sets were processed and analyzed with the NMRPipe program[14] and signal intensities were quantified using the program MUNIN[15]. Relaxation dispersion data were analyzed using a two-state exchange model, $A \underset{k_{\mathrm{BA}}}{\overset{k_{\mathrm{AB}}}{\rightleftharpoons}} B$ and the best-fit model parameters were extracted as described previously[16] using the program CATIA[16] which is available from http://pound.med.utoronto.ca/software.html. Briefly, the parameters are obtained by minimization of the target function,

$$\chi^2(\xi) = \sum \frac{\left(R_{2,eff}^{clc}(\xi) - R_{2,eff}^{exp}\right)^2}{\left(\Delta R_{2,eff}^{exp}\right)^2} \qquad \text{(S2)}$$

where $R_{2,eff}^{exp}$ and $\Delta R_{2,eff}^{exp}$ are experimental effective relaxation rates and their uncertainties, respectively, $R_{2,eff}^{clc}(\xi)$ are calculated relaxation rates obtained by numerical integration of the Bloch-McConnell equation[17], $\xi = \{x_1,\ldots,x_{npar}\}$ denotes the set of adjustable model parameters (including $\Delta\varpi$) and the summation in Eq S2 is over the number of experimental data points.

Quantitative ^3J measurements on L99A mutant of T4 Lysozyme

Three-bond scalar coupling constants $^3J(C^{\alpha},C^{\delta 1})$ were measured for the Ile residues of the L99A T4 lysozyme mutant using the pulse scheme published previously[18]

with a constant time delay for coupling evolution set to T = 58 ms. The data were processed and analyzed with the NMRPipe program and signal intensities were quantified using the program Sparky[19]. Subsequently, $^{3}J(C^{\alpha},C^{\delta1})$ couplings were calculated from the ratio of the intensities of the diagonal peak $I_d=(^{13}C_{methyl},^{13}C_{methyl},^{1}H_{methyl})$ and the cross peak $I_c=(^{13}C_{methyl},^{13}C_{alpha},^{1}H_{methyl})$ from

$$I_c/I_d = \tan^2(\pi\ ^{3}J\ T) \qquad \text{(S3)}$$

$^{3}J(C^{\alpha},C^{\delta1})$ scalar couplings and $\delta(C^{\delta1})$ chemical shifts:

The following data were used to generate Figure 2 and Eq 3.

Protein L [20]

Residue	$^{3}J(C^{\alpha},C^{\delta1})$ (Hz)	$\delta(C^{\delta1})$ (ppm)
11	3.7±0.1	14.30
60	3.0±0.1	12.32

Human ubiquitin (measured in-house)

Residue	$^{3}J(C^{\alpha},C^{\delta1})$ (Hz)	$\delta(C^{\delta1})$ (ppm)
3	3.6	14.4
13	3.2	14.6
23	2.3	9.5
30	3.6	15.4
36	3.3	13.8
44	2.4	12.7
61	3.5	14.5

C-terminal SH2 domain of phospholipase C-γ1 (peptide free)[21]

Residue	$^{3}J(C^{\alpha},C^{\delta1})$ (Hz)	$\delta(C^{\delta1})$ (ppm)
47	3.39	13.8
55	2.84	13.1
81	1.90	8.7
99	2.42	11.2

T4 Lysozyme L99A mutant (see above)

Residue	$^{3}J(C^{\alpha},C^{\delta1})$ (Hz)	$\delta(C^{\delta1})$ (ppm)
3	3.32	15.08
9	3.93	15.74
17	3.74	14.65
27	3.58	16.49
29	3.44	13.50

50	2.05	9.36
58	3.86	14.50
78	3.80	13.72
100	3.28	13.44
150	3.63	14.55

Calmodulin (Cam-M13)[18] (bmrb no. 5770)

Residue	$^3J(C^{\alpha},C^{\delta1})$ (Hz)	$\delta(C^{\delta1})$ (ppm)
9	3.0±0.2	13.1
27	1.7±0.2	15.6
52	2.1±0.2	10.5
63	2.7±0.2	10.5
85	3.1±0.2	13.2
100	3.0±0.2	16.1
125	1.9±0.2	9.50
130	2.9±0.2	12.70

Rat Nedd4 WWIII domain - ENaC bP2 Peptide Complex[22] (bmrb no. 4963)

Residue	$^3J(C^{\alpha},C^{\delta1})$ (Hz)	$\delta(C^{\delta1})$ (ppm)
29	3.15±0.01	15.33
33	3.59±0.01	13.73
48	3.08±0.01	12.75

Maltose Binding Protein (bmrb no. 4354)[23]

Residue	$^3J(C^{\alpha},C^{\delta1})$ (Hz)	$\delta(C^{\delta1})$ (ppm)
9	3.03	14.50
33	1.93	10.25
59	3.67	13.81
60	3.38	14.63
79	2.70	12.33
104	1.81	14.65
108	1.40	10.05
116	3.43	13.72
132	3.20	12.40
161	3.27	13.14
199	3.32	14.23
226	3.48	11.81
266	3.49	13.82
317	2.35	12.54
329	2.47	12.34
333	2.78	12.77
368	3.74	15.28

Full author list for reference no. 3 of text.

(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 USA 2008*, **105**, 4685-4690.

References

(1) Neal, S.; Nip, A. M.; Zhang, H.; Wishart, D. S. *J Biomol NMR* **2003**, *26*, 215-40.
(2) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; Gaussian, Inc., Wallingford, CT: 2003.
(3) Hehre, W.; Ditchfie, R.; Pople, J. *J Chem Phys* **1972**, *56*, 2257.
(4) Ditchfie, R.; Hehre, W.; Pople, J. *J Chem Phys* **1971**, *54*, 724.
(5) Barone, V. In *Recent Advances in Density Functional Methods*; Chong, D. P., Ed.; World Scientific Publ. Co: Singapore, 1996; Vol. I.
(6) Cheeseman, J.; Trucks, G.; Keith, T.; Frisch, M. *J Chem Phys* **1996**, *104*, 5497-5509.
(7) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67-81.
(8) Chou, J. J.; Case, D. A.; Bax, A. *J Am Chem Soc* **2003**, *125*, 8959-66.
(9) Neudecker, P.; Zarrine-Afsar, A.; Choy, W.-Y.; Muhandiram, D. R.; Davidson, A. R.; Kay, L. E. *J Mol Biol* **2006**, *363*, 958-76.
(10) 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.
(11) Mulder, F. A.; Hon, B.; Mittermaier, A.; Dahlquist, F. W.; Kay, L. E. *J Am Chem Soc* **2002**, *124*, 1443-51.
(12) Lundstrom, P.; Vallurupalli, P.; Religa, T. L.; Dahlquist, F. W.; Kay, L. E. *J Biomol NMR* **2007**, *38*, 79-88.
(13) Tollinger, M.; Skrynnikov, N. R.; Mulder, F. A. A.; Forman-Kay, J. D.; Kay, L. E. *J Am Chem Soc* **2001**, *123*, 11341-11352.
(14) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277-293.
(15) Korzhnev, D.; Ibraghimov, I.; Billeter, M.; Orekhov, V. *J Biomol NMR* **2001**, *21*, 263-268.
(16) Hansen, D. F.; Vallurupalli, P.; Lundstrom, P.; Neudecker, P.; Kay, L. E. *J Am Chem Soc* **2008**, *130*, 2667-75.
(17) McConnell, H. M. *J. Chem. Phys.* **1958**, *28*, 430-431.
(18) Bax, A.; Max, D.; Zax, D. *J. Am. Chem. Soc.* **1992**, *114*, 6923-6925.
(19) Kneller, D. G.; Kuntz, I. D. *J. Cell. Biochem.* **1993**, 254-254.

(20) Millet, O.; Mittermaier, A.; Baker, D.; Kay, L. E. *J. Mol. Biol.* **2003**, *329*, 551-563.
(21) Kay, L. E.; Muhandiram, D. R.; Farrow, N. A.; Aubin, Y.; Forman-Kay, J. D. *Biochemistry* **1996**, *35*, 361-368.
(22) Kanelis, V.; Rotin, D.; Forman-Kay, J. D. *Nat Struct Biol* **2001**, *8*, 407-12.
(23) Gardner, K. H.; Zhang, X.; Gehring, K.; Kay, L. E. *J. Am. Chem. Soc.* **1998**, *120*, 11738-11748.