Observation of CH···π Interactions between Methyl and Carbonyl Groups in Proteins.

Frédéric A Perras, Dominique Marion, Jérôme Boisbouvier, David L Bryce, Michael J Plevin

Protein Structure

International Edition: DOI: 10.1002/anie.201702626
German Edition: DOI: 10.1002/ange.201702626

Observation of CH···π Interactions between Methyl and Carbonyl Groups in Proteins

Frédéric A. Perras, Dominique Marion, Jérôme Boisbouvier, David L. Bryce,* and Michael J. Plevin**

Abstract: *Protein structure and function is dependent on myriad noncovalent interactions. Direct detection and characterization of these weak interactions in large biomolecules, such as proteins, is experimentally challenging. Herein, we report the first observation and measurement of long-range "through-space" scalar couplings between methyl and backbone carbonyl groups in proteins. These J couplings are indicative of the presence of noncovalent C–H···π hydrogen-bond-like interactions involving the amide π network. Experimentally detected scalar couplings were corroborated by a natural bond orbital analysis, which revealed the orbital nature of the interaction and the origins of the through-space J couplings. The experimental observation of this type of CH···π interaction adds a new dimension to the study of protein structure, function, and dynamics by NMR spectroscopy.*

The structure, stability, and function of proteins and other biomacromolecules are dependent on the formation and disruption of a variety of noncovalent interactions. The hydrogen bond is a well-studied noncovalent interaction that is prevalent in biochemistry. In proteins, the canonical hydrogen bonds that occur between backbone amide and carbonyl groups underpin the formation of α-helix and β-sheet secondary structure elements. In the last two decades, the importance of weaker varieties of the hydrogen bond has grown.[1–3] The current definition of a hydrogen bond includes interactions that involve soft acid and/or base moieties,[4] such as C–H···A (A = N, O, or S), D–H···π (D = N or O), and C–H···π interactions. These weaker interactions have been shown to exist in nature: C_α–H···O=C hydrogen bonds are often observed in β-sheet structures in proteins,[5,6] whereas C–H···O=C hydrogen bonds involving aliphatic donor groups have been reported in informatics surveys of protein 3D structures.[7,8] C–H···π interactions, in which the donor group interacts with the π orbitals of a conjugated moiety, are also known to be a common structural feature, with up to 15% of aromatic groups in proteins participating in this type of interaction.[9–11] Similarly to aromatic groups, the amide N–C=O network also forms a π-conjugated system that can participate in weak noncovalent interactions. For example, n→π* interactions involving electron donation from the lone pair of an oxygen atom to the unoccupied antibonding π* orbital of a carbonyl carbon atom have a strong presence in proteins.[12,13]

Noncovalent electronic interactions can be conveniently studied by characterizing the *J* coupling interaction, wherein nuclear spin–spin coupling is mediated by intervening electrons.[14] For example, canonical hydrogen bonds,[15] C_α–H···O=C hydrogen bonds,[6] C–H···π interactions,[11] and even van der Waals interactions[16] have been shown to lead to measurable *J* coupling interactions, thus bringing the electronic significance of these interactions onto a stronger experimental footing. Herein, on the basis of density functional theory (DFT) and solution NMR spectroscopy, we demonstrate that the backbone N–C=O network can act as an acceptor in CH···π interactions in proteins.

DFT calculations on a model ethane–glycinamide heterodimer predict the existence of *J* couplings on the order of tens of millihertz between carbon and proton nuclei involved in C_{methyl}–H···π_{CO} hydrogen-bond-like interactions (Figure 1). The magnitude of the ${}^{n_\pi}J_{HMeCO}$ coupling is dependent on the distance, *d*, between the methyl carbon atom and the center of the C=O bond. The coupling constant increases to a maximum at a distance of approximately 3.7 Å before decreasing at longer distances.

DFT-calculated ${}^{n_\pi}J_{HMeCO}$ coupling constants are fairly insensitive to the angle that the C–H bond forms with the carbonyl group (i.e., the angle φ), but are quite sensitive to the orientation with respect to the amide plane (i.e., the angle θ). A larger *J* coupling constant is calculated when the methyl group is oriented away from the normal of the amide plane, which is somewhat surprising given that one would expect that a C–H···π interaction would be strongest with the CH bond oriented directly above the carbonyl π-bonding orbital, as observed for other C–H···π interactions.[3,11] This difference is probably caused by a larger contribution from a C–H···O=C hydrogen-bond-type interaction.

[*] Dr. F. A. Perras, Prof. Dr. D. L. Bryce
Department of Chemistry and Biomolecular Sciences
University of Ottawa
10 Marie Curie Street, Ottawa, Ontario (Canada)
E-mail: david.bryce@uottawa.ca
Dr. D. Marion, Dr. J. Boisbouvier
Université Grenoble Alpes, CNRS, CEAEA
Institut de Biologie Structurale (IBS)
71 Avenue des Martyrs, 38044, Grenoble (France)
E-mail: jerome.boisbouvier@ibs.fr
Dr. M. J. Plevin
Department of Biology, University of York
Wentworth Way, York, YO10 5DD (UK)
E-mail: michael.plevin@york.ac.uk
Dr. F. A. Perras
Current address: US Department of Energy, Ames Laboratory
213 Spedding Hall, Ames, IA 50011 (USA)
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201702626.

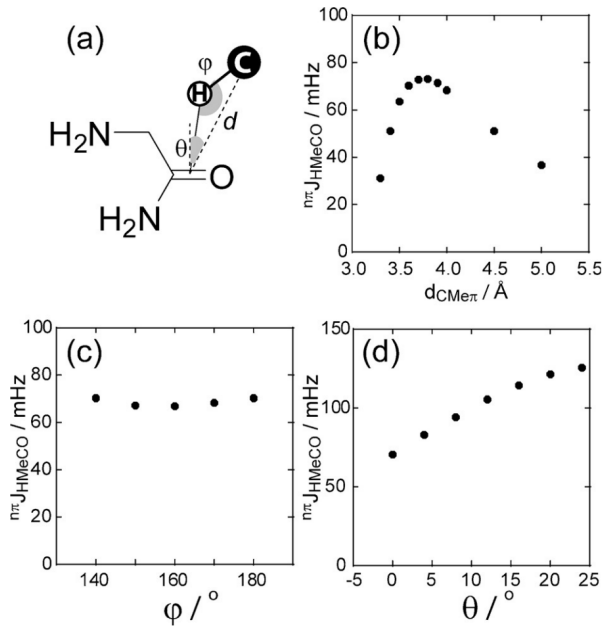

Figure 1. DFT analysis of $^{n_\pi}J_{HMeCO}$ coupling constants with a glycinamide–ethane model. a) Definitions of the distances and angles. b–d) Plots of the calculated $^{n_\pi}J_{HMeCO}$ coupling constants as a function of the C_{Me}–CO distance (b), the angle φ (c), and the angle θ (d). The default values of d, φ, and θ used for the calculations are 3.6 Å, 180°, and 0°, respectively. Rapid rotation of the methyl group was taken into consideration by averaging the coupling constants calculated for each of the three proton positions.

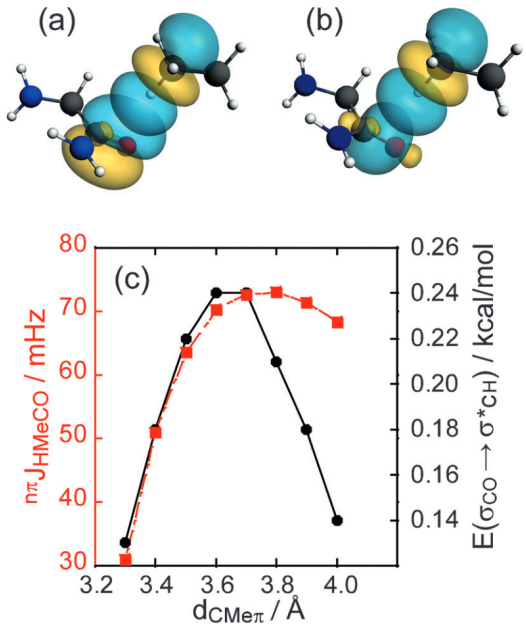

Figure 2. Results of NBO analysis. a, b) Orbitals corresponding to the dominant $\pi_{CO}\rightarrow\sigma^*_{CH}$ NBO interaction (a) and the $\sigma_{CO}\rightarrow\sigma^*_{CH}$ NBO interaction (b). c) Plot of the calculated $^{n_\pi}J_{HMeCO}$ coupling constants as a function of distance (red, primary y axis) alongside the $\sigma_{CO}\rightarrow\sigma^*_{CH}$ NBO interaction energy (black, secondary y axis).

To gain a greater insight into the origins of this noncovalent interaction, we carried out a natural bonding orbital (NBO) analysis.[17] The NBO interaction leading to the largest stabilizing effect is a $\pi_{CO}\rightarrow\sigma^*_{CH}$ interaction, which corresponds to a C–H···π hydrogen bond. Contributions from $\sigma_{CO}\rightarrow\sigma^*_{CH}$ and $\sigma_{CH}\rightarrow\pi^*_{CO}$ NBO interactions (Figure 2) also occur but are significantly weaker (see Table S1 and Figure S1 in the Supporting Information).

When the intermolecular $^{n_\pi}J_{HMeCO}$ coupling is decomposed into contributions arising from individual natural localized molecular orbitals (NLMOs),[18,19] the largest contributions arise from the σ_{CH} and σ_{CO} NLMOs, which would suggest that the $\sigma_{CO}\rightarrow\sigma^*_{CH}$ interaction contributes the most to the $^{n_\pi}J_{HMeCO}$ coupling, whereas the most energetically important contribution is the C–H···π hydrogen bond. These contributions are also evident when the J coupling and the energy of the $\sigma_{CO}\rightarrow\sigma^*_{CH}$ hydrogen-bond component are plotted as a function of the distance between the methyl and the carbonyl groups. The size of the $\sigma_{CO}\rightarrow\sigma^*_{CH}$ interaction is largest when the $^{n_\pi}J_{HMeCO}$ coupling constant is largest (Figure 2; see also Figure S1).

To experimentally characterize these weak interactions, we modified a 2D (^{1}H,^{13}C) heteronuclear multiple quantum coherence (HMQC) NMR experiment to enable the detection of weak long-range J couplings (see the Supporting Information). Given the small magnitude of the DFT-predicted $^{n_\pi}J_{HMeCO}$ coupling constants, the optimal transfer delay is governed by the transverse relaxation rates of the donor-group protons rather than by the magnitude of the coupling. To extend the transverse relaxation time (T_2) of these protons, we prepared two [U-^{2}H,^{13}C]-labeled samples of the protein ubiquitin with the specific incorporation of ($^{12}C^1H_3$) isotopomers in either Ile-δ_1 or Leu/Val-pro*S* methyl groups (see the Supporting Information for details on the isotopic labeling and pulse sequence). This isotopic labeling scheme preserves the favorable relaxation properties arising from intra methyl group interference effects[20] and enables the use of transfer delays of several hundred milliseconds without truncation effects due to efficient intra-residue transfer. To further increase sensitivity, we modified the standard NMR signal acquisition approach in such a way that ^{1}H detection was initiated before the end of the transfer delay (see Figure S2). Detection of both sides of the echo gives a theoretical improvement in the signal-to-noise ratio (S/N) of $\sqrt{2}$.

Long-range 2D ($^1H_{Me}$,$^{13}C_{CO}$) HMQC spectra revealed a number of cross-peaks that could only arise as a result of through-space $^{n_\pi}J_{HMeCO}$ coupling (Figure 3). To assign the observed C_{Me}-H···π_{CO} correlations, we recorded 3D HNCO NMR spectra with the same ubiquitin samples and identical spectral widths and resolution (Figure 4). $^{n_\pi}J_{HMeCO}$-coupled nuclei were assigned by comparing the measured ^{1}H and ^{13}C resonance frequencies with assigned resonance frequencies of methyl and carbonyl nuclei in ubiquitin. Internuclear distances extracted from the 3D structure of ubiquitin (PDB code: 1UBQ)[21] were used to exclude nuclei separated by more than 6.0 Å. In situations in which more than one carbonyl acceptor met chemical-shift and distance criteria, the group with the largest DFT-predicted $^{n_\pi}J_{HMeCO}$ coupling was selected (see Table S2). There are several candidate carbonyl acceptors for the peak at 175.0 ppm associated with

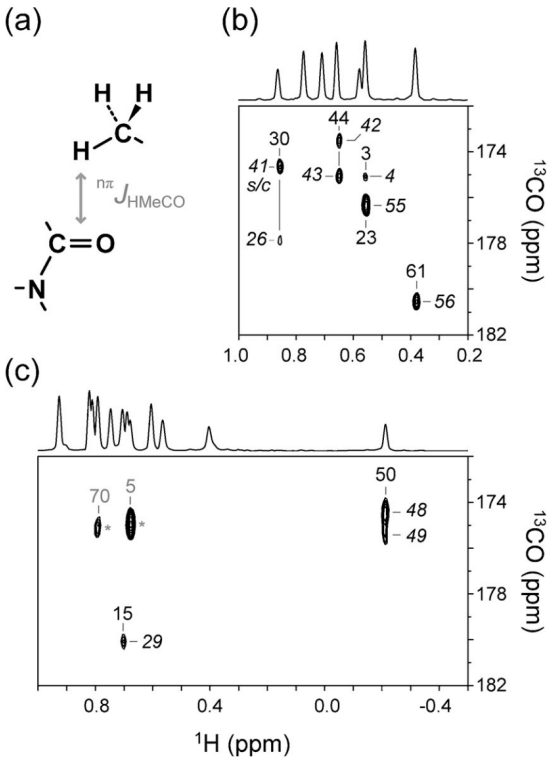

Figure 3. Detection of $^{n_\pi}J_{HMeCO}$ coupling in proteins. a) The through-space $^{n_\pi}J_{HMeCO}$ coupling detected between methyl protons and backbone carbonyl carbon atoms. b) Carbonyl region of a 2D (^{1}H,^{13}C) HMQC spectrum of [*U*-^{2}H,^{13}C,^{15}N], Ile-[δ_1-^{12}C^{1}H$_3$]-labeled ubiquitin. "s/c" indicates a side-chain carbonyl acceptor. c) Carbonyl region of a 2D (^{1}H,^{13}C) HMQC spectrum of [*U*-^{2}H,^{13}C,^{15}N], Leu/Val-[^{12}C^{1}H$_3$]pro*S*-labeled ubiquitin. Asterisks indicate cross-peaks arising from a through-bond ^5J coupling. Above each data set is the 1D projection of a 2D (^{1}H,^{13}C) SOFAST-HMQC spectrum of a [^{13}C^{1}H$_3$]methyl-labeled ubiquitin sample.

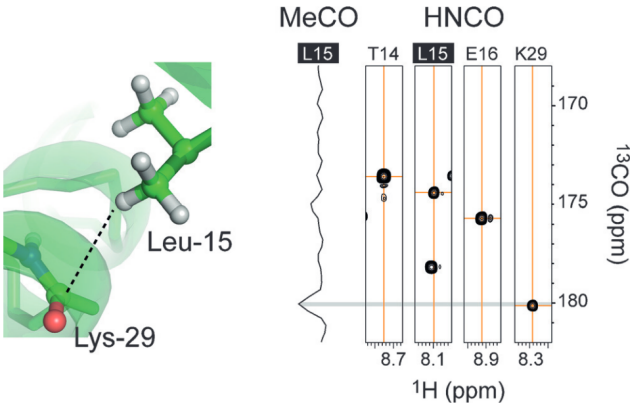

Figure 4. Assignment of Me/π_{CO} couplings. A 1D (^{13}C) trace of an "MeCO" spectrum of ubiquitin corresponding to the ^{1}H resonance frequency of the pro*S* methyl group of Leu-15 (0.72 ppm) is shown next to (^{1}H,^{13}C) strips extracted from a 3D (^{1}H,^{13}C,^{15}N) HNCO spectrum. HNCO strips were selected on the basis of sequence proximity to the donor residue ($i-1$, i, and $i+1$) or Me/CO distance ($d<6$ Å) and matching CO chemical shifts.

Ile-44. It was not possible to distinguish Leu-43 or Gln-49 by using DFT results alone. Of the two candidates, the carbonyl carbon atom of Leu-43 has the shortest internuclear distance ($d=2.8$ Å) and a more favorable geometry. To exclude the possibility that these signals arise from a weak through-bond ^6J coupling between Ile-44-δ_1 and Leu-43-CO nuclei, we conducted DFT calculations with a model in which the Ile-44 side chain was replaced with a methane molecule (situated at the δ_1 position). These analyses showed that a through-space $^{n_\pi}J_{HMeCO}$ coupling could yield the cross-peak seen in Figure 3 and that the carbonyl group Leu-43 was the most likely acceptor (see Table S2).

Two cross-peaks were observed at ^{1}H resonance frequencies corresponding to Val-Cγ_2 methyl-group protons. The corresponding ^{13}C chemical shifts indicated that these cross-peaks originated from long-range five-bond J couplings between a valine Cγ_2 donor group and the carbonyl group of the preceding residue (indicated by asterisks in Figure 3).

All DFT-calculated J coupling constants are positive with the exception of that for the Ile-61/Leu-56 pair. Since this spin pair features a geometry most similar to a C–H···O=C hydrogen bond (see Figures S4 and S5), the J coupling is expected to have the opposite sign, as the σ^*_{CH} NBO would interact with a lobe of opposite sign on the σ_{CO} NBO (see Figure S5), thus inverting the Fermi contact contribution to J.[19]

To evaluate the magnitude of the $^{n_\pi}J$ couplings associated with $C_{Me}H\cdots\pi_{CO}$ interactions, we acquired a data set with the ^{13}C carrier frequency set at 140 ppm and a ^{13}C spectral width of 200 ppm. In this experiment, magnetization can be transferred via $^{n_\pi}J$ couplings associated with CH/π interactions that involve either aromatic or carbonyl acceptor groups. The intensities of correlations that correspond to $C_{Me}H\cdots\pi_{CO}$ interactions are weaker than those corresponding to $C_{Me}H\cdots\pi_{aromatic}$ interactions (see Figure S6), as predicted from the DFT data presented herein. Moreover, the intensities of correlations arising from $C_{Me}H\cdots\pi_{CO}$ interactions involving Ile-δ_1 methyl-group donors and backbone carbonyl acceptors correlate well with the DFT-calculated scalar couplings (see Figure S6).

Analysis of the distribution of the $C_{Me}H\cdots\pi_{CO}$ interactions observed in ubiquitin revealed no apparent relationship between the detection of a C–H···π interaction and the carbonyl donor group being involved in a canonical hydrogen bond,[22] nor a preference for a particular type of secondary structure (see Figure S7). All CO acceptors identified in this study reside in peptide groups with average or above-average backbone order parameters (see Figure S7).[23] However, more variability in the dynamics of the methyl donor groups is evident (see Figure S8).[24] Notably, the δ_1 methyl group of Ile-44 has a lower than average order parameter, yet participates in at least two $C_{Me}H\cdots\pi_{CO}$ interactions. All methyl-group donors, and the majority of carbonyl-group acceptors, for which $C_{Me}H\cdots\pi_{CO}$ correlations were observed are inaccessible to the solvent (see Figure S8). The most interesting example in this context is the $C_{Me}H\cdots\pi_{CO}$ interaction involving the side-chain peptide group of Gln-41 (Figure 3). This residue is completely buried in an apolar environment, and the polar side-chain peptide group forms no canonical hydrogen bonds. Gln-41 is the only example of a buried asparagine or gluatamine residue in ubiquitin.

The detection of small through-space $^{n_\pi}J_{HMeCO}$ couplings provides strong experimental support for the presence of C–H···π interactions involving peptide-bond acceptor groups in proteins. The labeling scheme employed herein only permitted the detection of $C_{Me}H\cdots\pi_{CO}$ interactions involving Ile-δ_1 and Leu-pro*S* methyl groups. However, C–H···π_{CO} interactions in proteins are by no means limited to these two methyl groups. $C_{Me}H\cdots\pi_{CO}$ interactions involving other methyl donor groups could be similarly investigated by NMR spectroscopy by using alternative methyl-labeling approaches,[25] whereas non-methyl aliphatic C–H donor groups could be characterized by using approaches such as stereoarray isotope labeling.[26] Moreover, other chemical groups (e.g. hydroxy groups) in proteins could also form X-H···π_{CO} interactions and be investigated by NMR spectroscopy by the use of an appropriately isotopically labeled protein sample.

Even though these noncovalent interactions are weaker than canonical hydrogen bonds, C–H···π_{CO} interactions are present in large numbers in proteins[3] and are therefore likely to make an important cumulative contribution to the structure, dynamics, and function of a protein. Exploitation of the small and conformationally sensitive *J* couplings associated with weak noncovalent interactions involving π-acceptor groups promises to be a valuable tool for interrogating the contribution of these hydrogen-bond-like interactions to the structure and function of proteins.

Acknowledgements

F.A.P. thanks NSERC for a graduate scholarship and a Banting Postdoctoral Fellowship. F.A.P. is also supported through a Spedding Fellowship funded by the Laboratory Directed Research and Development (LDRD) program at the Ames Laboratory. The Ames Laboratory is operated for the US Department of Energy by Iowa State University under Contract No. DE-AC02-07CH11358. The high-field NMR and isotopic-labeling facilities at the Grenoble Instruct Centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) were used for this research with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). Research leading to these results was funded by the European Research Council under the EU Seventh Framework Program FP7/2007–2013 Grant Agreement no. 260887. D.L.B. thanks NSERC for funding. We thank Dr. Remy Sounier for assistance with sample preparation and Dr. Gail Bartlett for critical reading of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Keywords: C–H···π interactions · density functional calculations · *J* coupling · NMR spectroscopy · proteins

How to cite: *Angew. Chem. Int. Ed.* **2017**, *56*, 7564–7567
Angew. Chem. **2017**, *129*, 7672–7675

[1] M. S. Weiss, M. Brandl, J. Suhnel, D. Pal, R. Hilgenfeld, *Trends Biochem. Sci.* **2001**, *26*, 521–523.
[2] M. Nishio, *Phys. Chem. Chem. Phys.* **2011**, *13*, 13873–13900.
[3] M. Nishio, Y. Umezawa, J. Fantini, M. S. Weiss, P. Chakrabarti, *Phys. Chem. Chem. Phys.* **2014**, *16*, 12648–12683.
[4] G. R. Desiraju, *Angew. Chem. Int. Ed.* **2011**, *50*, 52–59; *Angew. Chem.* **2011**, *123*, 52–60.
[5] Z. S. Derewenda, L. Lee, U. Derewenda, *J. Mol. Biol.* **1995**, *252*, 248–262.
[6] F. Cordier, M. Barfield, S. Grzesiek, *J. Am. Chem. Soc.* **2003**, *125*, 15750–15751.
[7] J. D. Yesselman, S. Horowitz, C. L. Brooks, R. C. Trievel, *Proteins Struct. Funct. Bioinf.* **2015**, *83*, 403–410.
[8] G. J. Bartlett, D. N. Woolfson, *Protein Sci.* **2016**, *25*, 887–897.
[9] T. Steiner, G. Koellner, *J. Mol. Biol.* **2001**, *305*, 535–557.
[10] M. Brandl, M. S. Weiss, A. Jabs, J. Suhnel, R. Hilgenfeld, *J. Mol. Biol.* **2001**, *307*, 357–377.
[11] M. J. Plevin, D. L. Bryce, J. Boisbouvier, *Nat. Chem.* **2010**, *2*, 466–471.
[12] G. J. Bartlett, A. Choudhary, R. T. Raines, D. N. Woolfson, *Nat. Chem. Biol.* **2010**, *6*, 615–620.
[13] A. Choudhary, R. T. Raines, *Protein Sci.* **2011**, *20*, 1077–1081.
[14] J.-C. Hierso, *Chem. Rev.* **2014**, *114*, 4838–4867.
[15] G. Cornilescu, J.-S. Hu, A. Bax, *J. Am. Chem. Soc.* **1999**, *121*, 2949–2950.
[16] M. P. Ledbetter, G. Saielli, A. Bagno, N. Tran, M. V. Romalis, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12393–12397.
[17] J. P. Foster, F. Weinhold, *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.
[18] A. E. Reed, F. Weinhold, *J. Chem. Phys.* **1985**, *83*, 1736–1740.
[19] J. Autschbach, B. Le Guennic, *J. Chem. Educ.* **2007**, *84*, 156.
[20] V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, L. E. Kay, *J. Am. Chem. Soc.* **2003**, *125*, 10420–10428.
[21] S. Vijay-Kumar, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* **1987**, *194*, 531–544.
[22] S. Grzesiek, F. Cordier, V. Jaravine, M. Barfield, *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, *45*, 275–300.
[23] N. Tjandra, S. E. Feller, R. W. Pastor, A. Bax, *J. Am. Chem. Soc.* **1995**, *117*, 12562–12566.
[24] C. Farès, N.-A. Lakomek, K. F. A. Walter, B. T. C. Frank, J. Meiler, S. Becker, C. Griesinger, *J. Biomol. NMR* **2009**, *45*, 23–44.
[25] R. Kerfah, M. J. Plevin, R. Sounier, P. Gans, J. Boisbouvier, *Curr. Opin. Struct. Biol.* **2015**, *32*, 113–122.
[26] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono, P. Guntert, *Nature* **2006**, *440*, 52–57.

Manuscript received: March 13, 2017
Version of record online: May 22, 2017

