How to cite: *Angew. Chem. Int. Ed.* **2024**, *63*, e202414324
doi.org/10.1002/anie.202414324

NMR Spectroscopy

Homonuclear Super-Resolution NMR Spectroscopy

*Olivia Gampp[+], Luca Wenchel[+], Peter Güntert, and Roland Riek**

Abstract: In homonuclear ^{1}H NMR (nuclear magnetic resonance) spectra such as [^{1}H,^{1}H]-NOESY (Nuclear Overhauser Enhancement spectroscopy), which is a historic cornerstone spectrum for biomolecular NMR structural biology, hundreds to thousands of cross peaks are present within a square of approximately 100 ppm^2 leading to a lot of signal overlap. Spectral resolution is thus a limiting factor for unambiguous chemical shift assignment and data interpretation for dynamics and structure elucidation. Acquiring the spectra at higher magnetic fields such as at a 1.2 GHz ^{1}H frequency helps to reduce spectral crowding, since resolution scales proportionally to the magnetic field strength. Here, we show that the linewidths of cross peaks in [^{1}H,^{1}H]-NOESY and [^{1}H,^{1}H]-TOCSY spectra can be further reduced by a factor of 2–3 in each dimension by super-resolution spectroscopy. In the indirect dimension a composite exponential-cosine weighted number of scans along the time increments are recorded and digitally smoothened by a window function, while in the direct dimension an exponential-cosine window function is applied. Furthermore, measurement time saving by reduced-acquisition super-resolution (RASR) is introduced. Application to the 20 kDa protein KRAS shows that highly resolved NMR spectra suitable for automated analysis can be acquired within less than 3 hours. The method opens an avenue towards automated chemical shift assignment, dynamics and structure determination of unlabeled small and medium size proteins within 24 hours.

[*] O. Gampp,[+] L. Wenchel,[+] Prof. P. Güntert, Prof. R. Riek
Institute of Molecular Physical Science, ETH Zürich, Vladimir-Prelog-Weg 2, CH-8093 Zürich, Switzerland
E-mail: roland.riek@phys.chem.ethz.ch

Prof. P. Güntert
Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt am Main, 60438 Frankfurt am Main, Germany

Prof. P. Güntert
Department of Chemistry, Tokyo Metropolitan University, 192-0397 Hachioji, Tokyo, Japan

[+] These authors contributed equally to this work.

With the advent of 2D (two-dimensional) homonuclear NMR spectroscopy, such as the [^{1}H,^{1}H]-NOESY (Nuclear Overhauser Enhancement Spectroscopy)[1–2] and [^{1}H,^{1}H]-TOCSY (total correlation spectroscopy)[3–4] experiments, chemical shift assignment and structure determination of small proteins and RNA/DNA was established.[5] The limitation to small proteins is due to severe cross peak overlap in these 2D spectra where hundreds to thousands of cross peaks are located within a small area of approximately 10×10 ppm. This limitation can be resolved by stable isotope labeling,[6–9] the extension to 3D experiments,[10] the introduction of about a dozen complementary triple resonance experiments,[11–12] and relaxation-optimized spectroscopy (i.e. TROSY).[13–15] However, this comes at the cost that the acquisition of these spectra takes several weeks, and the even longer task of assigning the signals manually must also be taken into account. If only the backbone assignment is of interest, a plethora of sophisticated pulse sequences have been proposed for optimizing both measurement and human analysis time.[16–17] Recently, the software package ARTINA established not only a fully automated procedure for chemical shift assignment and structure determination,[18] but also indicated that a limited set of 2D and 3D NMR spectra with a total experimental time of approximately 1 week is needed to assign the backbone.[19] With our recent introduction of super-resolution NMR spectroscopy,[20–21] a resolution enhancement of up to a factor of 2 can be achieved in hetero-nuclear correlation experiments ([^{15}N,^{1}H]-HSQC). Here, we present super-resolution [^{1}H,^{1}H]-NOESY and [^{1}H,^{1}H]-TOCSY experiments and spectra recorded at 1.2 GHz for the 20 kDa cancer-relevant protein KRAS.[22] The ultra-high resolution of these spectra opens the possibility for protein structure determination of unlabeled proteins within 24 h comprising both measurement and analysis time.

Super-resolution [^{1}H,^{1}H]-NOESY (with a mixing time of 100 ms) and [^{1}H,^{1}H]-TOCSY experiments (with a mixing time of 60 ms) of the protein KRAS (at a concentration of 1.3 mM in a 3 mm NMR tube), were acquired on a Bruker Avance NEO at 1.2 GHz ^{1}H frequency (28.2 T) equipped with a 3 mm TCI Cryoprobe at 298 K with TopSpin version 4.3.0 (Bruker). Processing and visualization of all NMR spectra were done with TopSpin or PROSA.[23] The principle of super-resolution spectroscopy in the indirect dimension is to record an exponentially increasing number of scans per increment (proportional to e^{Rt} with given R) to counteract the exponential decay proportional to $e^{-R_2 t}$ of the signal due to transverse relaxation with rate R_2.[20] This approach is called dynamic number of scan sampling (DNS) below. To optimize the signal-to-noise per measurement time (sensitiv-

ity), the exponential weighting is expanded to an exponential-cosine weighted measurement Scheme with the number of scans proportional to $e^{Rt}\cos(\pi t/(2t_{max}))$, in which the cosine component is equivalent to a conventionally applied cosine apodization function.[20,24] For the 20 kDa protein KRAS the exponential weighting function determining the number of scans for each time increment has an exponential weighting factor of $R=125$ Hz during the indirect time evolution of 1256×60 µs = 75.36 ms (with 1256 complex points and a spectral width of 1/(60 µs)). The acquisition time in the direct dimension was approximately doubled to 184 ms = 3072×60 µs (see also Table S1).

Figure 1 shows the high quality of the super-resolution [^{1}H,^{1}H]-NOESY spectrum of KRAS acquired for 37 h, corresponding to an average of 32 scans per time increment. The combination of the ultra-high field and the super-resolution method along the indirect dimension yields a spectrum with almost no apparent remaining peak overlap and many empty areas between cross peaks highlighting the amazing resolution (Figure 1a). The effect of the super-resolution is evident when compared with a conventional

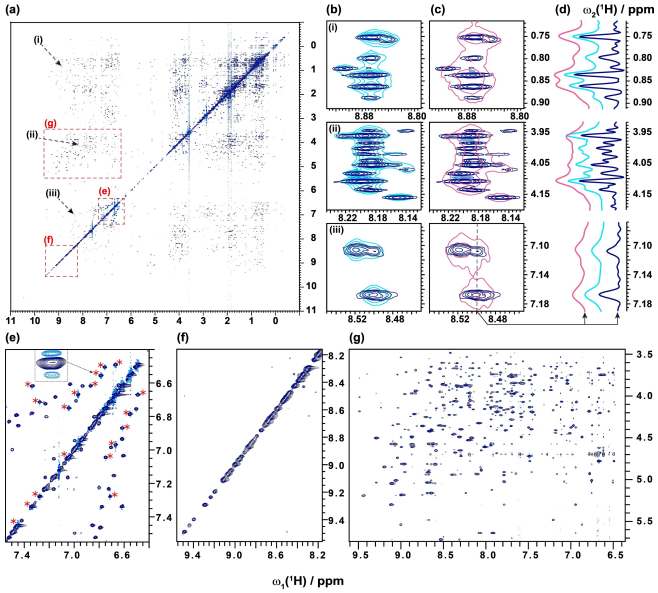

Figure 1. Super-resolution [^{1}H,^{1}H]-NOESY of the unlabeled KRAS compared to conventional acquisition. a) SR-[^{1}H,^{1}H]-NOESY of the KRAS measured by super-resolution using an exponential-cosine weighted dynamic number of scans and processed with the correction window function. The spectrum was measured with $R=125$ Hz, which corresponds approximately to half the transverse relaxation rate R_2. The spectrum was acquired at 1200 MHz ^{1}H frequency. b) Three enlarged segments of the SR-[^{1}H,^{1}H]-NOESY (blue) of KRAS overlayed with spectra acquired conventionally at 1200 MHz (cyan) or c) 700 MHz (pink). d) Cross sections from the segments in (b) and (c). e) Enlargement of the diagonal in the aromatic region of the spectrum. Cross-peaks marked by red asterisks are from Asn/Gln-protons, for which the SR effect was exaggerated because of their smaller R_2 relaxation rates, causing signal truncation yielding line shapes with wiggles. f) Enlargement of the diagonal in the H^N–H^N region of the spectrum showing the super-resolution effect. g) Enlargement of the H^α–H^N region of the spectrum showing how little signal overlap is present due to the high resolution.

[^{1}H,^{1}H]-NOESY spectrum measured at 1.2 GHz (superposition in Figure 1b). A factor of 2–3 cross peak sharpening was obtained as also evidenced by cross sections (Figure 1d). The combined effect of both the high field of 1.2 GHz and the super-resolution technique is shown in Figure 1c, where a corresponding standard [^{1}H,^{1}H]-NOESY spectrum measured at 700 MHz ^{1}H frequency is presented on top of the corresponding super-resolution 1.2 GHz spectrum.

The high spectral quality close to the diagonal could only be obtained by digital smoothing the discrete values of the exponential-cosine weighting function[23]. The discrete nature is due to the discrete number of scans and the STATES-TPPI acquisition method, which requires a four-step phase cycle, and therefore the number of scans for each time increment to be an integer multiple of 4. This discrete step recording NS_{disc} can be corrected to the desired continuous $NS_{cont} \infty e^{Rt}\cos(\pi t/(2t_{max}))$ as shown in Figure S1c by multiplying the signal by NS_{cont}/NS_{disc}. While this smoothing was not considered important for the hetero-nuclear experiments,[20] it is essential for the homonuclear experiments where the diagonal peaks are strong relative to the cross peaks of interest. Figure S1 shows two spectra obtained from the same measured time-domain data with (Figure S1c) or without (Figure S1a) smoothing, demonstrating the importance of the smoothing function. As an alternative to smoothing, the starting number of scans can be increased as shown in Figure S1a versus S1b, where, without smoothing, a [^{1}H,^{1}H]-NOESY of KRAS acquired with 4 scans in the initial increment is compared with the corresponding spectrum acquired with 32 initial scans. While the minimal step size for the number of scans is 4 in both cases, the artifacts of the latter spectrum are much smaller (while still present), since the step size is smaller in proportion to the total number of scans for each time increment.

Other artifacts visible in the NMR spectrum are due to inhomogeneous transverse ^{1}H relaxation. While using an exponential factor with $R=125$ Hz the cross-peak linewidth due to transverse relaxation for the backbone amide and aliphatic hydrogens was reduced as anticipated yielding sharp cross peaks, the effect was exaggerated for signals from slowly relaxing entities such as for example Asn/Gln side chain moieties causing truncation artifacts manifested by line shapes with wiggles as highlighted in Figure 1e by an asterisk and in Figure S3.

The super-resolution concept in the indirect dimension can also be applied digitally,[20] meaning that a regular [^{1}H,^{1}H]-NOESY spectrum is measured with a constant number of scans per increment (for example, 4) but with the same high maximal acquisition time as (i.e. 75.36 ms) followed by the multiplication of a composite exponential-cosine weighted window function (EC WF) as demonstrated in Figure 2. However, as predicted from theoretical considerations and demonstrated by cross sections in Figures 2 and S2 the exponential-cosine weighted function approach is about 20 % less sensitive than the dynamic number of scan sampling (DNS) super-resolution measurement approach discussed above. It was shown analytically that the DNS recording of super-resolution data is always superior to the

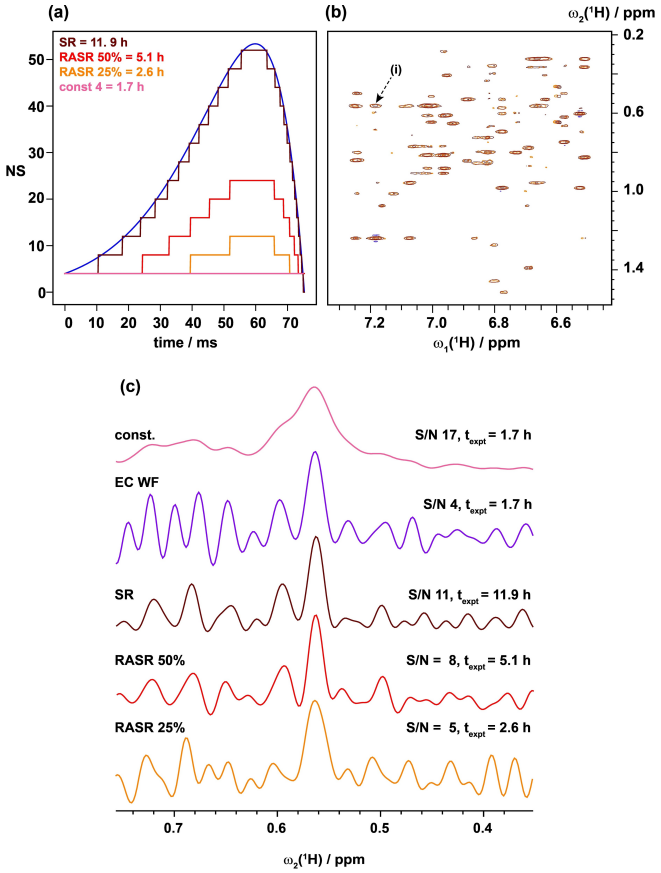

Figure 2. Super-resolution (SR) and reduced-acquired super-resolution (RASR) [^{1}H,^{1}H]-NOESY of unlabeled KRAS. a) Acquired number of scans (NS) as a function of the indirect evolution time t_1 displayed for super-resolution with 100% of dynamic number of scans acquired in burgundy (SR), super-resolution with a 50% of the dynamic number of scans acquired scheme (RASR 50%) in red, super-resolution with 25% of the dynamic number of scans acquired scheme (RASR 25%) in yellow, and conventional 4-scan acquisition in pink. The exponential-cosine weighted window function, to which SR and RASR schemes are smoothed, is plotted in dark blue. b) A region of the SR-[^{1}H,^{1}H]-NOESY of the KRAS measured by super-resolution using an exponential-cosine weighted dynamic number of scans and processed with the correction window function (burgundy) is overlayed with the SR-[^{1}H,^{1}H]-NOESY RASR acquired with only 25% of the scans and processed with the correction window function in yellow. The two spectra are very similar, and no peak shifts are observed. c) Cross sections along the indirect dimension of the peak indicated by (i) in (b) using the color coding in (a) for the different acquisition methods. The exponential-cosine weighted window function applied to the conventionally acquired spectrum (const.) is given in purple (EC WF). The signal-to-noise ratio and the total measurement time are given for each acquisition method.

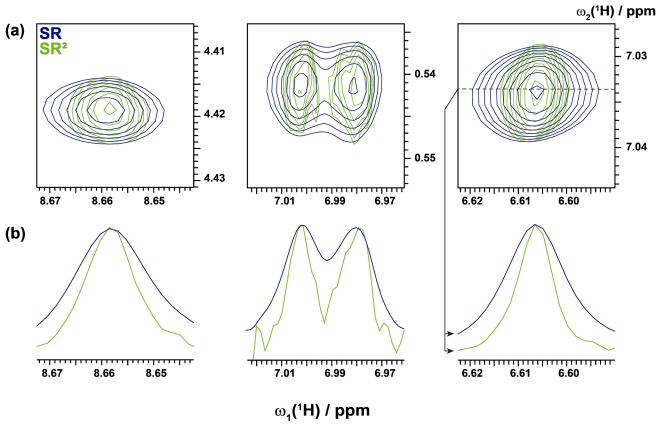

Figure 3. Super-resolution [^{1}H,^{1}H]-NOESY of the unlabeled KRAS with further resolution enhancement in the direct dimension using an exponential growing cosine-modulated window function yielding super-resolution in both dimensions, SR×SR = SR^2. a) Three segments of the SR-[^{1}H,^{1}H]-NOESY of KRAS measured by super-resolution using an exponential-cosine weighted dynamic number of scans and processed with the correction window function (blue; same spectrum as in Figure 1) are overlayed with the spectrum processed with an exponential-cosine window function in the direct (acquisition) dimension with R = 125 Hz (green), which corresponds approximately to half the transverse relaxation rate R_2. The same value of R was used for the super-resolution measurement in the indirect dimension. The spectrum was acquired at 1200 MHz ^{1}H frequency. b) Cross sections from the segments in (a) along the direct ^{1}H dimension.

multiplication of the FID with a corresponding window function.[20] An exponential-cosine weighting window function (EC WF) is however of great interest in the direct acquisition dimension as shown in Figure 3 yielding also cross peak sharpening in the order of a factor of two along the direct dimension (termed SR^2 = SR * SR in both dimensions).

Finally, to decrease and optimize the measurement time the reduced acquisition super-resolution (RASR) method is introduced, which allows for the reduction of the measurement time from 12 hours in the super-resolution method to 2.6 hours for the KRAS [^{1}H,^{1}H]-NOESY without the drawbacks of prediction tools such as non-uniform sampling (NUS). With RASR a composite DNS/exponential-cosine weighted window function is selected, which, for example, combines acquiring 50% of the dynamic number of scans, which is smoothed by a window function to reproduce the full dynamic sampling of Figure 1 as shown in Figure 3. Similarly, only 25% of the dynamic number of scan sampling can be smoothed to the function to yield a super-resolution spectrum with the effect that the measurement time is reduced by a factor of 4 (Figure 2) along with little signal loss (Figure 2c). It is also superior in terms of signal-to-noise compared to applying only the exponential-cosine weighted window function (Figure 2c), which is expected from theory. It is noted that with RASR every time increment is sampled and thus there is no shifting of signals or appearance of spurious peaks which may occur with NUS or machine learning-based spectral reconstruction techniques.[25–26] This is the case because processing the RASR spectra requires no reconstruction program and is only based on a slightly modified version of the super-resolution smoothing window function which does not touch the oscillating frequencies. Additionally, there is no need to optimize sampling Schemes. The number of scans acquired along the indirect dimension follows the distribution of the exponential-cosine weighted function. The RASR procedure is considered of particular importance for the [^{1}H,^{1}H]-NOESY spectrum, where cross peaks belonging to a specific

resonance must be aligned precisely to the same frequency within the spectrum for proper assignment, which is not guaranteed by non-Fourier reconstruction methods.

The fast acquisition of super-resolved [^{1}H,^{1}H]-NOESY spectra recorded at 1.2 GHz for a 20 kDa protein with thousands of resolved cross peaks and resonances lacking significant signal overlap suggests the possibility of fully automated chemical shift assignment and structure determination within a day including both measurements and analysis. Since it is evident that similar resolution can be obtained for [^{1}H,^{1}H]-TOCSY spectra relevant for the sequential assignment as shown for KRAS in Figure 4, deep learning-based automation for assignment and structure calculation using 2D NMR spectroscopy may be feasible and is currently in progress.

This is supported by theoretical considerations in the Supporting Information that indicate how the expected overlap in NMR spectra scales as a function of protein size, spectrometer frequency, spectrum dimensionality, and other linewidth-affecting factors (such as use of the present super-resolution method). These considerations show that the same peak overlap in a 2D homonuclear spectrum is expected for a protein with double the molecular mass when the magnetic field is doubled (for example from 600 MHz to 1.2 GHz ^{1}H frequency) and super-resolution with a factor of two resolution enhancement is applied (Table S2). Furthermore, because of practical constraints on the measurement of indirect dimensions in 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectra, where for small and medium size proteins the linewidth is limited by the number of sampled time points rather than by transverse relaxation, 2D [^{1}H,^{1}H]-NOESY will exhibit only slightly less resolved peaks than the 3D version recorded under same conditions. As shown in Table S1, if 2D [^{1}H,^{1}H]-NOESY spectra have either 50, 80, or 90 % resolved peaks, the corresponding 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectra are expected to have 58, 84, or 92 % resolved peaks, respectively (Table S1), a modest improvement considering the increased measurement time.

It is fascinating to realize that 40 years after establishing structure determination of proteins by liquid state NMR three key ingredients—the 1.2 GHz NMR spectrometer equipped with a cryo-probe, super-resolution spectroscopy,[27] and the advent of machine learning-based methods to replace tedious manual homonuclear NOESY spectrum analysis and assignment—suggest going back to the initial 2D experiments available.

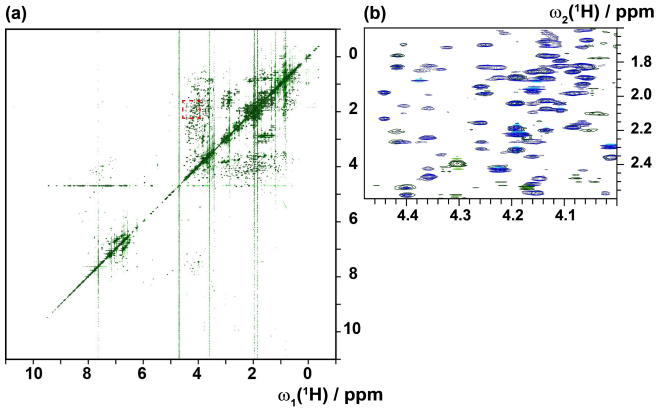

Figure 4. Super-resolution [^{1}H,^{1}H]-TOCSY spectrum of unlabeled KRAS. a) SR-CLEAN [^{1}H,^{1}H]-TOCSY measured with super-resolution using an exponential-cosine weighted dynamic number of scans and processed with a digital smoothing function. b) Close-up of the region marked in red in (a) shown in green overlayed with the corresponding region in the SR-[^{1}H,^{1}H]-NOESY spectrum (blue).

Supporting Information

The authors have cited additional references within the Supporting Information.[28–30]

Acknowledgements

We thank ETH Zurich and the Swiss National Science Foundation for financial support. Open Access funding provided by Eidgenössische Technische Hochschule Zürich.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: NMR · protein · NOESY · linewidth · super-resolution

[1] A. Kumar, R. R. Ernst, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
[2] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546–4553.
[3] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, *53*, 521–528.
[4] C. Griesinger, G. Otting, K. Wüthrich, R. R. Ernst, *J. Am. Chem. Soc.* **1988**, *110*, 7870–7872.
[5] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, **1986**.
[6] A. Bax, M. A. Weiss, *J. Magn. Reson.* **1987**, *71*, 571–575.
[7] L. P. McIntosh, F. W. Dahlquist, *Q. Rev. Biophys.* **1990**, *23*, 1–38.
[8] R. L. Isaacson, P. J. Simpson, M. Liu, E. Cota, X. Zhang, P. Freemont, S. Matthews, *J. Am. Chem. Soc.* **2007**, *129*, 15428–15429.
[9] V. Tugarinov, L. E. Kay, *J. Biomol. NMR* **2004**, *28*, 165–172.
[10] H. Oschkinat, C. Griesinger, P. J. Kraulis, O. W. Sørensen, R. R. Ernst, A. M. Gronenborn, G. M. Clore, *Nature* **1988**, *332*, 374–376.
[11] L. E. Kay, M. Ikura, R. Tschudin, A. Bax, *J. Magn. Reson.* **1990**, *89*, 496–514.

[12] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, N. J. Skelton, M. Rance, *Protein NMR Spectroscopy. Principles and Practice*, Academic Press, San Diego, CA, **2007**.
[13] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *J. Am. Chem. Soc.* **1998**, *120*, 6394–6400.
[14] R. Riek, K. Pervushin, K. Wüthrich, *Trends Biochem. Sci.* **2000**, *25*, 462–468.
[15] A. G. Tzakos, C. R. R. Grace, P. J. Lukavsky, R. Riek, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 319–342.
[16] P. W. Coote, S. A. Robson, A. Dubey, A. Boeszoermenyi, M. X. Zhao, G. Wagner, H. Arthanari, *Nat. Commun.* **2018**, *9*, 3014.
[17] C. R. R. Grace, R. Riek, *J. Am. Chem. Soc.* **2003**, *125*, 16104–16113.
[18] P. Klukowski, R. Riek, P. Güntert, *Nat. Commun.* **2022**, *13*, 6151.
[19] P. Klukowski, R. Riek, P. Güntert, *Sci. Adv.* **2023**, *9*, eadi9323.
[20] L. Wenchel, O. Gampp, R. Riek, *J. Magn. Reson.* **2024**, *366*, 107746.
[21] C. Waudby, J. Christodoulou, *ChemRxiv preprint* **2020**, DOI: 10.26434/chemrxiv.12006750.v1.
[22] J. L. Bos, E. R. Fearon, S. R. Hamilton, M. Verlaandevries, J. H. Vanboom, A. J. Vandereb, B. Vogelstein, *Nature* **1987**, *327*, 293–297.
[23] P. Güntert, V. Dötsch, G. Wider, K. Wüthrich, *J. Biomol. NMR* **1992**, *2*, 619–629.
[24] B. Simon, H. Köstler, *J. Biomol. NMR* **2019**, *73*, 155–165.
[25] M. Mobli, J. C. Hoch, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 21–41.
[26] S. G. Hyberts, H. Arthanari, G. Wagner, *Top. Curr. Chem.* **2012**, *316*, 125–148.
[27] C. F. DeLisle, H. B. Mendis, J. L. Lorieau, *J. Biomol. NMR* **2019**, *73*, 105–116.
[28] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. M. Ono, P. Güntert, *Nature* **2006**, *440*, 52–57.
[29] F. Hefke, R. Schmucki, P. Güntert, *J. Biomol. NMR* **2013**, *56*, 113–123.
[30] G. S. Rule, *Fundamentals of Protein NMR Spectroscopy*, Springer, New York, **2006**.

Manuscript received: June 29, 2024
Accepted manuscript online: September 30, 2024
Version of record online: November 7, 2024