DOI: 10.1002/cbic.201200044

Fast Automated NMR Spectroscopy of Short-Lived Biological Samples

Suhas Tikole,[a] Victor Jaravine,[a] Vladimir V. Rogov,[a, b] Alexis Rozenknop,[a, c] Kerstin Schmöe,[a] Frank Löhr,[a] Volker Dötsch,[a] and Peter Güntert*[a, d, e]

Protein NMR structure and interaction studies are often complicated by sample-related issues such as instability and proteolytic degradation,[1] as well as spectral factors: poor signal dispersion, overlap, dynamic exchange, and weak resonances. In solution NMR spectroscopy, the stability of proteins is significantly affected by the aqueous conditions.[2] Under realistic physiological conditions, protein samples often aggregate and precipitate with concomitant loss of activity.[3] This can severely limit the quality of the measurements, or even render systems impossible to study. Here we address the challenge of developing methodology to overcome these issues by enabling fast but comprehensive NMR measurements before a sample degrades.

Optimal sample conditions from the point of view of experimental NMR spectroscopy are often at odds with what is best for the sample itself. For example, to improve NMR sensitivity, high temperature may be selected as it reduces solvent viscosity. Similarly denaturation or destabilization narrows signals, and less salt is preferred because the presence of salt reduces considerably the sensitivity of normal and cold probes.[4] Thus, it is not uncommon for NMR spectroscopists to choose sample conditions that are far from the native ones. If the sample is unstable under measurement conditions, it can be necessary to prepare multiple samples, and to replace them in the course of recording the NMR spectra upon sample degradation or loss.

Determining NMR protein structure relies on almost complete and accurate backbone and side-chain resonance assignments, which in turn rely on the completeness of multidimensional $^{1}H,^{13}C,^{15}N$-triple resonance datasets.[5] Longer experiments generally yield higher quality information. Multidimensional NMR experiments often require long measurement times, especially if high resolution is necessary to resolve spectral overlap for large or unfolded proteins.[6] Currently, a state-of-the-art set of a dozen complete experiments for a protein assignment and structure determination requires about one month of measurement time.[7]

While corresponding sample lifetimes are often achievable for purified proteins in aqueous solution, this is too long for most samples in their native environment, such as in-cell or cell-lysate preparations. If many spectra are needed for the backbone and side-chain chemical shift assignment and the structure calculation, it becomes impossible to achieve high-resolution in all experiments in realistic time. For short-lived proteins, this usually results in a severely compromised resolution that is up to an order of magnitude less than what is possible for a similar long-lived protein. Another common compromise is concentration versus time, as the sensitivity of signals depends linearly on the protein concentration, but higher concentrations can result in faster multimerization and precipitation. This is aggravated by short sample lifetimes. Decay occurring during the experiments can result in corrupted data and the loss of essential information for assignments and structure calculations. The problem becomes more severe for in-cell or unpurified cell-lysate preparations as they contain a cocktail of biologically active ingredients.

Here we present a new strategy for obtaining high-quality information for complete resonance assignments of short-lived target proteins in native cell lysate by using a "universal" fusion system with a small leader protein, a modified ubiquitin (see the preceding paper by Dötsch et al.).[8] Native cell lysates allow realistic functional studies while avoiding precipitation and similar problems during isolation/purification. We suggest recording a "maximal dataset" with short total measurement time by optimizing the sensitivity and resolution of fast NMR spectroscopy. As a proof of principle, we applied the method to the 23 kDa protein RcsD-ABL-HPt for identifying the function of the RcsD-ABL domain in the Rcs phosphorelay,[9] and to several short-lived, ubiquitin-fused peptides (see Section 1 in the Supporting Information). We measured sets of non-uniformly sampled (NUS)[10] experiments (Section 2 in the Supporting Information) and performed peak picking (Section 3 in the Supporting Information) and automated resonance assignments[11] (Section 4 in the Supporting Information) of isolated and purified Ub2_NBR1-LIR and Ub2_p62-LIR constructs[12] as test examples and of the short-lived Ub2_p62-LIR and Ub3_NBR1-LIR constructs in cell lysate.[8]

The RcsD-ABL-HPt construct possesses residues 688–890 of the *E. coli* RcsD histidine kinase and contains the two inde-

[a] *S. Tikole, Dr. V. Jaravine, Dr. V. V. Rogov, Dr. A. Rozenknop, Dr. K. Schmöe, Dr. F. Löhr, Prof. Dr. V. Dötsch, Prof. Dr. P. Güntert*
Institute of Biophysical Chemistry and
Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt
Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)
E-mail: guentert@em.uni-frankfurt.de

[b] *Dr. V. V. Rogov*
Institute of Protein Research
142290 Pushchino (Russia)

[c] *Dr. A. Rozenknop*
Institute of Biochemistry II and
Cluster of Excellence Macromolecular Complexes
Goethe University Frankfurt
Max-von-Laue Strasse 7, 60438 Frankfurt am Main (Germany)

[d] *Prof. Dr. P. Güntert*
Frankfurt Institute for Advanced Studies
Goethe University Frankfurt
Ruth-Moufang-Strasse 1, 60438 Frankfurt am Main (Germany)

[e] *Prof. Dr. P. Güntert*
Graduate School of Science, Tokyo Metropolitan University
1-1 Minami-ohsawa, Hachioji, Tokyo 192-0397 (Japan)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201200044.

pendent ABL and HPt domains.[9] The NMR spectra of RcsD-ABL-HPt displayed low signal dispersion due to a significant content of flexible loops and α-helices, and were thus suitable for showing the performance of the method on a relatively large system requiring high-resolution data.

Such samples are interesting for functional studies in cell lysate but undergo rapid degradation. This is illustrated by the disappearance of target peptide resonances of Ub2_p62-LIR (Figure 1). Note that the chemical shifts are similar in water

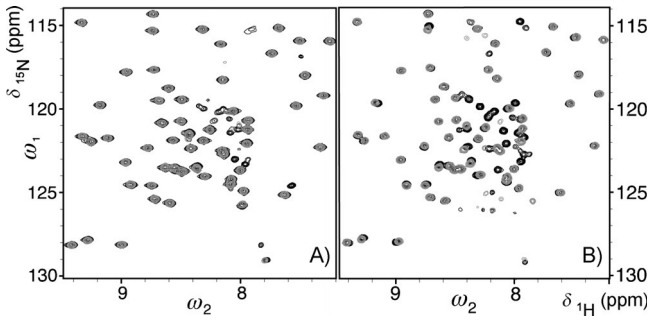

Figure 1. ^{1}H,^{15}N HSQC spectra showing the degradation of target peptides. A) Purified Ub2_p62-LIR. B) Ub2_p62-LIR cell lysate. Peaks shown in black were observed only in freshly prepared samples. Those shown in gray remain visible after three weeks in the purified sample (A), or after three days in the cell lysate sample (B).

and in cell lysate, although these solution environments are very different.

Accurate NMR measurements and chemical-shift assignments are important for protein-interaction studies. A small perturbation in the chemical shifts might indicate a different conformation or interaction. This includes both the backbone and side-chain resonances. Although there are many backbone-assignment programs,[13] obtaining side-chain assignments automatically is still a challenge. Few such programs exist, and most side-chain assignments are still done manually or sometimes semimanually. The FLYA protocol,[11c] which makes use of the GARANT algorithm for automated chemical-shift assignment[11a,b] was shown to be able to obtain complete chemical-shift assignments. For this protocol to work well, increased spectral resolution is important.

NUS provides high spectral resolution,[14] while hyperdimensional (HD) NMR spectroscopy enables NUS with much smaller fractions of the complete data matrices than if each experiment were processed separately.[15] Consequently, more experiments can be recorded in the same allocated measurement time, for example, ten 10% NUS experiments can be recorded in the same time as one fully sampled one. This opens up the possibility of working with "maximal sets", rather than the usual approach of minimal sets, when the measurement time is limited. In contrast to approaches that require specialized pulse sequences and fixed datasets, any standard experiment can be utilized in our method.

The use of HD NMR spectroscopy with NUS was demonstrated on proteins of various sizes for backbone resonance assignment[16] and in real time.[17] The HD method derives both direct and indirect correlations between various spin systems. To adapt this method for rapid measurements of NMR signals in short-lived samples and to include side-chain assignment, we optimized NUS–HD spectral processing, signal analysis, and automated assignment procedures for a larger experimental set and for differing targets. Our aim was to develop an approach that is generally applicable to the structural characterization of long- and short-lived large and small proteins as well as non-folded peptides, and has filtering and sensitivity enhancement capabilities.

The assignment results for Ub3_NBR1-LIR in cell lysate (Figure 2), purified Ub2_NBR1-LIR (Figure S5 in the Supporting Information), and RcsD-ABL-HPt (Table S2) show that increasing

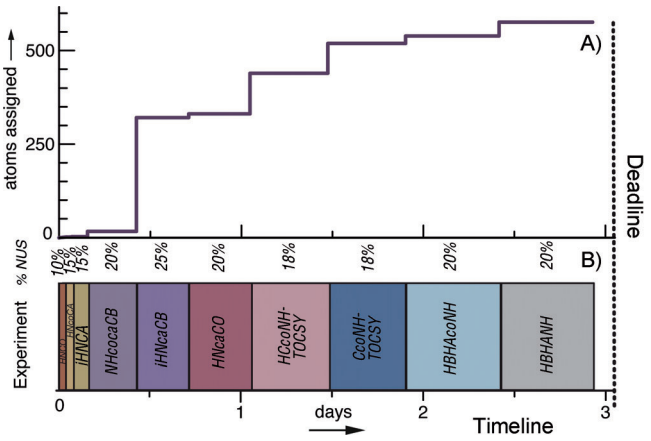

Figure 2. Timeline of recording the NMR experiments for the "cell-lysate" sample Ub3_NBR1-LIR, and assignment of its backbone and side-chain resonances. A) Number of consolidated[11c] resonance assignments. B) Experiments in the dataset. The width of each rectangle is proportional to the duration of the experiment. The percentage numbers on top are sparse fractions. Since all these experiments are essentially needed for the assignment procedure, the experiments had to be recorded with a fraction of full grid sampling in order to record the complete dataset before the expected deadline of the sample lifetime of around three days. The experiments add cumulatively to the assignment. The order of experiments is "sorted", that is, the first recorded ones are most sensitive with shorter durations.

the number of experiments in the set leads to an increase in the number of assigned chemical shifts for all systems, as expected. An improvement in the assignment was also observed as a result of simultaneous HD processing[16] compared to treating each NUS spectrum individually (Figure 3). To optimize the chemical-shift assignment, we varied the number of automatically picked peaks and recalculated the assignments for each relative peak-picking (RPP) factor (Figure 3).

RPP is defined as the number of picked peaks relative to the number of peaks expected for the protein sequence[11b,18] (see Section 3 in the Supporting Information). As a function of the RPP, the extent of correct assignments for the fused Ub2_NBR1-LIR protein including both the peptide target and ubiquitin increased from close to zero to a maximum value (Figure 3). The assignment quality decreases slowly for RPP factors above 0.8 as gradually higher number of picked noise features interfered with the assignment. The most complete

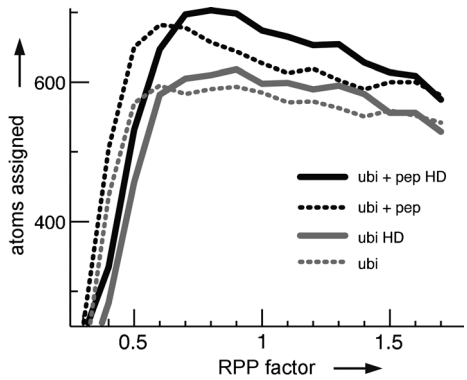

Figure 3. Comparison of the resonance assignment of Ub2_NBR1-LIR with (solid lines) and without HD processing (dashed lines). The numbers of consolidated assigned atoms are given as a function of the RPP factor for the entire fused Ub2_NBR1-LIR protein (black) and the ubiquitin part alone (gray).

assignment was obtained with an RPP factor of 0.8 instead of the theoretically expected value of 1.0 because the flexible loop connecting ubiquitin with the peptide target and comprising ~16% of the sequence was not observed for the fused protein, and there were nearly no assignments for the loop region residues 77–95 for all RPP factors. Figure 4 illustrates the backbone and side-chain chemical shift assignments for the ubiquitin and the target peptide part.

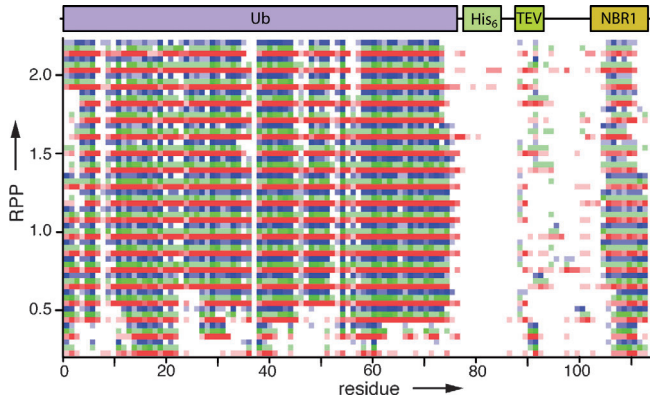

Figure 4. Backbone and side-chain assignments for the fused protein Ub2_NBR1-LIR calculated by using the FLYA algorithm and plotted as a function of the RPP scaling factor. Assignments are represented by colored rectangles: backbone atoms in red, side-chain carbon atoms in green, and side-chain hydrogen atoms in blue. The color intensity is proportional to the percentage of "consolidated" assigned atoms in the corresponding category of atoms. The domains and sequence segments are shown at the top.

Figure S2 shows the signal-to-noise (S/N) ratio for four Ub2_NBR1-LIR peak lists as a function of the RPP factor. For RPP values below 1.1–1.5, HD processing improves the S/N by about 30% compared to non-HD processing. For higher RPPs, the average S/N ratios degrade as a large number of noisy features are picked up. Abundant noise peaks at high RPP levels can introduce errors into the assignment, whereas improved S/N ratios of the peak lists translate into more reliable assignments. Finding the optimal number of peaked picks is thus important for the automated assignment.

In the spectra of cell lysate samples, the signals of the protein of interest can interfere with those of other proteins[19] and cellular components; this can introduce additional artifacts into the NMR spectra, although these artifacts can be eliminated by HD processing. Only the signals selected in the reference spectrum are retained in the HD spectra if all spectra are processed simultaneously with a single reference spectrum. Apart from filtering, this improves the average S/N ratio for all spectra compared to individually processed spectra (Figure S1 and Table S1). Thus, HD processing enhances the sensitivity of the target protein signals.

As the fused constructs enhanced the protein yield, relatively high sample concentrations enabled NMR studies directly in the cell lysate without any isolation and purification procedures. Thus the combination of improved fast NMR spectroscopy methods and the increased peptide life time (described in the preceding paper by Dötsch et al.[8]) provide the possibility of measuring all the spectra necessary for backbone and side-chain resonance assignment in three days by utilizing 10–30 mL of $^{13}C,^{15}N$-labeled M9-based cell culture. Varying the sparse fractions allows for S/N optimization within the allocated total measurement time, when time savings from the most sensitive experiments can be used either for longer recording of the less-sensitive ones or to record additional experiments. The time can be shortened further if higher protein concentration can be achieved.

To summarize, we have developed a fast NMR spectroscopy method by using optimized NUS–HD processing to characterize peptides and proteins that are known to have a short life span. The method was initially tested on several purified protein samples, including a 23 kDa, two-domain protein and several fused peptides, and then applied to study samples in native cell lysate without isolation or purification procedures. The results show that automated resonance assignment and fast characterizations of the structure, interactions, and functions of short-lived proteins are feasible.

Experimental Section

Details of the recording and processing of spectra, peak picking, and automated resonance assignment are given in Sections 1–4 of the Supporting Information. For instance, ten 10–30% NUS experiments both for backbone and side-chain resonance assignment were recorded for the purified Ub2_NBR1-LIR protein with a total measurement time of 4.6 days, whereas for a cell lysate sample of the same protein (Ub3_NBR1-LIR), a very similar set of ten experiments was recorded in 2.9 days (Table S1). Full sampling of these sets would have taken 26.5 or 15.2 days, respectively. A further reduction in the measurement time for the cell lysate sample was achieved by utilizing BEST pulse sequences.[20] Resonance assignments were performed automatically by using the FLYA algorithm[11c] implemented in CYANA.[21]

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (grant DFG JA1952/1–1), the Lichtenberg program of the Volkswagen Foundation, and the Japan Society for the Promotion of Science (JSPS). We thank Natalia Rogova for help with the sample preparations, Elena Schmidt and Sina Kazemi for advice with the FLYA calculations, and Wolfgang Bermel (Bruker GmbH) for help with the pulse sequences and NUS software implementations on the spectrometer.

Keywords: cell lysates · hyperdimensional processing · NMR spectroscopy · peak-picking · proteins · resonance assignment

[1] G. T. Montelione, D. Y. Zheng, Y. P. J. Huang, K. C. Gunsalus, T. Szyperski, *Nat. Struct. Biol.* **2000**, *7*, 982 – 985.
[2] A. Yee, A. Gutmanas, C. H. Arrowsmith, *Curr. Opin. Struct. Biol.* **2006**, *16*, 611 – 617.
[3] a) J. L. Cleland, M. F. Powell, S. J. Shire, *Crit. Rev. Ther. Drug Carrier Syst.* **1993**, *10*, 307 – 377; b) W. Wang, *Int. J. Pharm.* **2005**, *289*, 1 – 30; c) S. E. Bondos, A. Bicknell, *Anal. Biochem.* **2003**, *316*, 223 – 231.
[4] P. F. Flynn, D. L. Mattiello, H. D. W. Hill, A. J. Wand, *J. Am. Chem. Soc.* **2000**, *122*, 4823 – 4824.
[5] a) K. H. Gardner, L. E. Kay, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 357 – 406; b) G. Wider, K. Wüthrich, *Curr. Opin. Struct. Biol.* **1999**, *9*, 594 – 601; c) D. Y. Zheng, Y. P. J. Huang, H. N. B. Moseley, R. Xiao, J. Aramini, G. V. T. Swapna, G. T. Montelione, *Protein Sci.* **2003**, *12*, 1232 – 1246.
[6] a) M. Hennig, W. Bermel, A. Spencer, C. M. Dobson, L. J. Smith, H. Schwalbe, *J. Mol. Biol.* **1999**, *288*, 705 – 723; b) J. Yao, H. J. Dyson, P. E. Wright, *FEBS Lett.* **1997**, *419*, 285 – 289; c) O. Zhang, J. D. Forman-Kay, D. Shortle, L. E. Kay, *J. Biomol. NMR* **1997**, *9*, 181 – 200.
[7] H. S. Atreya, T. Szyperski, *Methods Enzymol.* **2005**, *394*, 78 – 108.
[8] V. V. Rogov, A. Rozenknop, N. Y. Rogova, F. Löhr, S. Tikole, V. Jaravine, P. Güntert, I. Dikic, V. Dötsch, *ChemBioChem*, **2012**; DOI: 10.1002/cbic.201200045.
[9] K. Schmöe, V. V. Rogov, N. Y. Rogova, F. Löhr, P. Güntert, F. Bernhard, V. Dötsch, *Structure* **2011**, *19*, 577 – 587.
[10] J. C. Hoch, A. S. Stern, *NMR Data Processing*, Wiley, New York, **1996**.
[11] a) C. Bartels, M. Billeter, P. Güntert, K. Wüthrich, *J. Biomol. NMR* **1996**, *7*, 207 – 213; b) C. Bartels, P. Güntert, M. Billeter, K. Wüthrich, *J. Comput. Chem.* **1997**, *18*, 139 – 149; c) B. Lopez-Mendez, P. Güntert, *J. Am. Chem. Soc.* **2006**, *128*, 13112 – 13122.
[12] A. Rozenknop, V. V. Rogov, N. Y. Rogova, F. Löhr, P. Güntert, I. Dikic, V. Dötsch, *J. Mol. Biol.* **2011**, *410*, 477 – 487.
[13] a) J. A. Lukin, A. P. Gove, S. N. Talukdar, C. Ho, *J. Biomol. NMR* **1997**, *9*, 151 – 166; b) D. E. Zimmerman, C. A. Kulikowski, Y. P. Huang, W. Q. Feng, M. Tashiro, S. Shimotakahara, C. Y. Chien, R. Powers, G. T. Montelione, *J. Mol. Biol.* **1997**, *269*, 592 – 610; c) N. E. G. Buchler, E. R. P. Zuiderweg, H. Wang, R. A. Goldstein, *J. Magn. Reson.* **1997**, *125*, 34 – 42; d) M. Leutner, R. M. Gschwind, J. Liermann, C. Schwarz, G. Gemmecker, H. Kessler, *J. Biomol. NMR* **1998**, *11*, 31 – 43; e) K. B. Li, B. C. Sanctuary, *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 467 – 477.
[14] D. Rovnyak, J. C. Hoch, A. S. Stern, G. Wagner, *J. Biomol. NMR* **2004**, *30*, 1 – 10.
[15] a) E. Kupce, R. Freeman, *J. Am. Chem. Soc.* **2006**, *128*, 6020 – 6021; b) E. Kupce, R. Freeman, *Prog. Nucl. Magn. Reson. Spectrosc.* **2008**, *52*, 22 – 30; c) E. Lescop, B. Brutscher, *J. Am. Chem. Soc.* **2007**, *129*, 11916 – 11917.
[16] V. A. Jaravine, A. V. Zhuravleva, P. Permi, I. Ibraghimov, V. Y. Orekhov, *J. Am. Chem. Soc.* **2008**, *130*, 3927 – 3936.
[17] V. A. Jaravine, V. Y. Orekhov, *J. Am. Chem. Soc.* **2006**, *128*, 13421 – 13426.
[18] R. Schmucki, S. Yokoyama, P. Güntert, *J. Biomol. NMR* **2009**, *43*, 97 – 109.
[19] Z. Serber, A. T. Keatinge-Clay, R. Ledwidge, A. E. Kelly, S. M. Miller, V. Dötsch, *J. Am. Chem. Soc.* **2001**, *123*, 2446 – 2447.
[20] A. Favier, B. Brutscher, *J. Biomol. NMR* **2011**, *49*, 9 – 15.
[21] P. Güntert, *Eur. Biophys. J.* **2009**, *38*, 129 – 143.

Received: January 18, 2012
Published online on April 11, 2012

