

Part 1
In-cell NMR Spectroscopy:
From 2001 to Now

CHAPTER 1

In-cell NMR Spectroscopy: From 2001 to Now

VOLKER DÖTSCH

Institute of Biophysical Chemistry, Goethe University, Frankfurt, Germany
Email: vdoetsch@em.uni-frankfurt.de

1.1 In-cell NMR – The Basis

The dream of cell biologists is to observe their macromolecule of interest directly in its natural environment, inside a cell. Fluorescence microscopy has revolutionized the way proteins and nucleic acids can be visualized within this complicated, crowded surrounding. The advantage of fluorescence microscopy is its very high sensitivity and, in combination with super-resolution techniques, the ability to obtain high spatial resolution. This high sensitivity also allows for a high temporal resolution as the time it takes to detect signals with reasonable signal-to-noise ratios is short. However, detecting the localization of macromolecules in the cellular environment is not enough to understand their biological function. Interaction with other cellular components can only be obtained *via* fluorescence correlation microscopy, requiring two fluorescently tagged molecules. Interactions with a non-fluorescent molecule or conformational changes cannot be detected or only indirectly be proposed or require immunoprecipitation experiments. This gap in information can be – in principle – filled by NMR spectroscopy. During the last four decades, NMR spectroscopy has developed into a very important analytical tool in the chemical and biological sciences. In addition to its ability to determine structures of biological macromolecules under

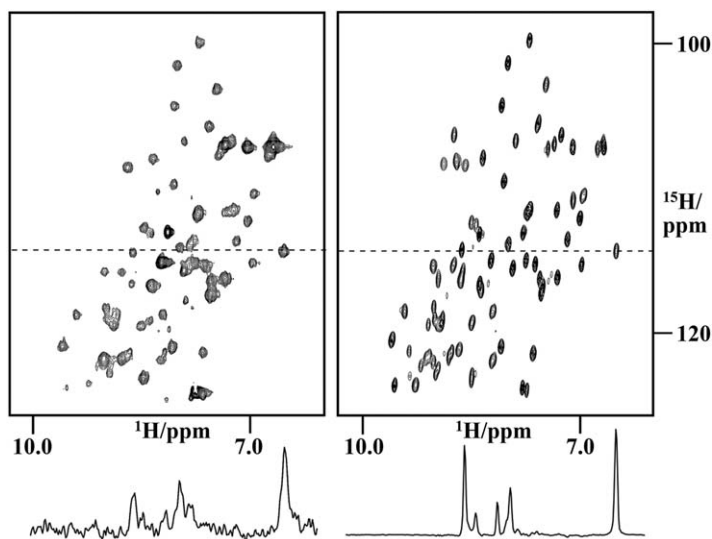


Figure 1.1 First in-cell NMR spectrum of NmerA in the cytoplasm of BL29 *E. coli* cells (left) and comparison with the spectrum of the purified protein (right).

near-physiological conditions, NMR spectroscopy is capable of investigating the interaction between biological macromolecules and other molecules, ranging from proteins and nucleic acids to small organic compounds.

The ability to investigate intermolecular interactions is based on the sensitivity of the chemical shift to changes in the chemical environment. In addition to direct interaction with other molecules, changes in the chemical shifts of a biological macromolecule can also indicate changes to its conformation, for example different folding states. The development of in-cell NMR experiments that enable NMR investigations of biological macromolecules in their natural environment has created new tools for the direct investigation of the interaction between biological macromolecules and other macromolecules or drugs in living cells (Figure 1.1).¹⁻⁷ In addition, these methods allow researchers to obtain information about the conformation and dynamics of macromolecules in their natural environment and to compare the *in vivo* behavior with data obtained from *in vitro* studies. A unique application of in-cell NMR is to study the folding and maturation of proteins directly following expression in the cellular environment.^{6,8,9}

1.2 Detecting the Molecule of Interest in the Crowded Environment of the Cell

The observation of biological macromolecules in cellular systems requires labeling schemes that enable the selective identification of these macromolecules in an environment that is crowded with other macromolecules as

well as with many different small molecules.^{1,2,10–12} To make the macromolecule of interest visible in this crowded intracellular environment, macromolecules have to be labeled, either with NMR active isotopes or with paramagnetic probes for in-cell EPR studies.^{13–17} In principle, two different ways are possible to achieve labeling in the intracellular environment. The first method is the expression of the macromolecule of interest inside cells. The original in-cell NMR experiments used this overexpression method with bacterial cells,^{1,10,12} however, in the meantime overexpression has also been achieved with other cell types including different eukaryotic cells (yeast,¹⁸ insect cells,¹⁹ mammalian cells⁶). Overexpression, particularly in bacteria, is convenient and uses very well established protocols, but it has the potential of producing a huge amount of background signals since not only the macromolecule of interest but all cellular components will be labeled with NMR-active isotopes. The original in-cell NMR experiments in bacteria surprisingly demonstrated that uniform ¹⁵N-labeling of the overexpressed protein is feasible and that only a minimal level of background signals is produced.^{1,10} The only requirements for observing an overexpressed ¹⁵N-labeled protein in in-cell NMR experiments are that its expression level reaches a certain threshold and that its rotational correlation time is sufficiently short.¹⁰ The last condition is met in principle for all macromolecules that can also be investigated by NMR *in vitro* and that do not interact with large cellular components since the intracellular viscosity is only a factor of two higher than the viscosity of water.^{20,21}

In contrast, uniform labeling of overexpressed macromolecules with ¹³C results in a high level of background signals that makes unambiguous identification of the overexpressed macromolecule impossible.²² One exception is signals with unique chemical shifts such as high field shifted methyl groups or anomeric sugar resonances.^{12,22} The difference in the level of background signals between ¹⁵N and ¹³C was attributed to the fact that C–H groups are far more abundant in small molecules in the cell than N–H groups, and that many amide protons in small molecules are not protected from solvent exchange, thus making their NMR signals undetectable due to exchange line broadening.¹⁰

To overcome this problem of high background signals, amino acid-type selective labeling schemes have been suggested and successfully applied.²² In particular in bacterial cells, the use of isotopically labeled amino acids for selective labeling schemes has to be considered carefully, as amino acids often represent precursors for the synthesis of other amino acid types, thus leading to isotopic scrambling. Amino acids that are at the end of a metabolic pathway have been shown to be useful for in-cell NMR experiments such as Lys, Arg or His for ¹⁵N labeling and Met and with some restriction Ala for ¹³C labeling. Another strategy is to use labeled precursors, for example α -ketobutyrate to incorporate labeled isoleucine into the protein of interest in bacteria.²³ Likewise, the use of fluorinated amino acids for labeling of overexpressed proteins has been successfully used and produced background-free spectra.^{24,25}

A different strategy that was proposed is to suppress the high background by selective expression of the protein of interest with the concomitant suppression of the expression of all other (host) proteins. This suppression of host gene expression can be achieved by two different techniques. The first is based on using the drug rifampicin, which inhibits the bacterial RNA polymerase.²⁶ The RNA polymerase of the bacteriophage T7, however, is not affected by this drug, thus allowing the selective overexpression of the protein of interest from a plasmid containing a T7 promoter in bacterial cells expressing the T7 RNA polymerase. The second approach is degradation of all mRNA molecules that contain ACA sequences by an mRNA interferase.²⁷ Changing all ACA base triplets in the DNA sequence of the protein of interest by utilizing other codons enables the selective translation of the mRNA of a particular protein while all other mRNAs that contain ACA triplets are degraded.

1.3 In-cell NMR with Eukaryotic Cellular Systems

While these initial bacteria-based in-cell NMR experiments were used to study different aspects of protein behavior, for example protein–protein interactions as well as protein interactions with small molecules, it was clear from the beginning that the method had to be applied to eukaryotic cell systems as well. For these systems, new methods to get isotopically labeled macromolecules into the cellular interior were developed. The first one used is based on micro-injecting a highly concentrated sample of a purified macromolecule labeled with NMR-active isotopes into the cell type of interest.^{28,29} This method, however, is restricted to cells that are sufficiently large for such a mechanical injection, in particular *Xenopus laevis* oocytes. For smaller cells, the use of arginine-tags or dendrite-based systems for bringing bacterially expressed and isotopically labeled proteins across the cellular membrane has successfully been demonstrated.³⁰ In addition, the use of pore-forming toxins that can be sealed again with the application of calcium solutions³¹ and the formation of pores with electroporation has been established.³² The advantage of these techniques is that the natural abundance of the NMR-active isotopes represents the only potential background signals.

A critical parameter for in-cell NMR experiments is that a certain concentration must be reached. As in the bacterial systems, this can also be achieved in eukaryotic systems using overexpression protocols. Yeast and insect cells are two cell types that are frequently used to overexpress proteins for use in structural biology applications. Consequently, protocols for their use in in-cell NMR experiments based on overexpression of macromolecules of interest have been adapted. However, also certain mammalian cell types such as HEK293T cells can produce enough proteins to be detectable in in-cell NMR experiments and investigations of proteins in these cells following overexpression have been described.⁶

Not only the number of cell types that are used for in-cell NMR studies has been significantly expanded over the past 15 years but also more types of biological macromolecules and different magnetic resonance methods have been used. In addition to globular proteins, intrinsically unstructured polypeptides,^{33,34} RNA molecules and DNA molecules^{35,36} have been objects of in-cell NMR studies. In addition to one-dimensional and two-dimensional liquid-state NMR experiments, three-dimensional experiments using non-linear sampling have been combined with in-cell NMR experiments to measure distance constraints for structure calculations.³⁷ Proteins that tumble too slow to be detected in in-cell NMR experiments could be detected with the use of solid-state NMR,³⁸ which also provides a method to investigate membrane proteins in natural membranes.^{39,40} The use of in-cell EPR experiments has expanded the applicability of in-cell magnetic resonance to the investigation of larger biological macromolecules that do not produce high-resolution in-cell NMR spectra.¹⁵ A significant expansion of the applicability has been the development of systems that allow the continuous exchange of growth media in the NMR tube, thereby providing fresh nutrients as well as oxygen and enabling longer experiments under stable conditions.^{8,41,42}

1.4 The Aims of In-cell NMR

Applications of in-cell MR methods range from protein–protein and protein–drug interactions to investigation of ion binding, posttranslational modifications, folding and maturation to full three-dimensional structure determination. As the low sensitivity of NMR in general requires relatively high intracellular concentrations, most of the time studying the interactions with natural macromolecular binding partners is not possible as they would have to reach the same concentration range as the protein of interest. Therefore, investigations of the intracellular behavior of the macromolecule of interest (conformations, dynamics, non-specific interactions, effects of molecular crowding) as well as processes in which the protein of interest is enzymatically coupled to another cellular component, can be and have been studied with these techniques. Examples of these applications in many different cell types focusing on different biological macromolecules with different magnetic resonance techniques are discussed in detail in the following chapters.

References

1. Z. Serber, A. T. Keatinge-Clay, R. Ledwidge, A. E. Kelly, S. M. Miller and V. Dötsch, *J. Am. Chem. Soc.*, 2001, **123**, 2446–2447.
2. Z. Serber and V. Dötsch, *Biochemistry*, 2001, **40**, 14317–14323.
3. D. S. Burz, K. Dutta, D. Cowburn and A. Shekhtman, *Nat. Methods*, 2006, **3**, 91–93.

4. M. M. Dedmon, C. N. Patel, G. B. Young and G. J. Pielak, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12681–12684.
5. J. A. Hubbard, L. K. MacLachlan, G. W. King, J. J. Jones and A. P. Fosberry, *Mol. Microbiol.*, 2003, **49**, 1191–1200.
6. L. Banci, L. Barbieri, I. Bertini, E. Luchinat, E. Secci, Y. Zhao and A. R. Aricescu, *Nat. Chem. Biol.*, 2013, **9**, 297–299.
7. L. Barbieri, E. Luchinat and L. Banci, *Sci. Rep.*, 2015, **5**, 14456.
8. A. Mochizuki, A. Saso, Q. Zhao, S. Kubo, N. Nishida and I. Shimada, *J. Am. Chem. Soc.*, 2018, **140**, 3784–3790.
9. J. Danielsson, K. Inomata, S. Murayama, H. Tochio, L. Lang, M. Shirakawa and M. Oliveberg, *J. Am. Chem. Soc.*, 2013, **135**, 10266–10269.
10. Z. Serber, R. Ledwidge, S. M. Miller and V. Dötsch, *J. Am. Chem. Soc.*, 2001, **123**, 8895–8901.
11. Z. Serber, P. Selenko, R. Hansel, S. Reckel, F. Lohr, J. E. Ferrell, Jr., G. Wagner and V. Dötsch, *Nat. Protoc.*, 2006, **1**, 2701–2709.
12. J.-M. Wieruszkeski, A. Bohin, J.-P. Bohin and G. Lippens, *J. Magn. Reson.*, 2001, **151**, 118–123.
13. R. Igarashi, T. Sakai, H. Hara, T. Tenno, T. Tanaka, H. Tochio and M. Shirakawa, *J. Am. Chem. Soc.*, 2010, **132**, 8228–8229.
14. M. Azarkh, O. Okle, V. Singh, I. T. Seemann, J. S. Hartig, D. R. Dietrich and M. Drescher, *ChemBioChem*, 2011, **12**, 1992–1995.
15. M. Azarkh, V. Singh, O. Okle, I. T. Seemann, D. R. Dietrich, J. S. Hartig and M. Drescher, *Nat. Protoc.*, 2013, **8**, 131–147.
16. M. Qi, A. Gross, G. Jeschke, A. Godt and M. Drescher, *J. Am. Chem. Soc.*, 2014, **136**, 15366–15378.
17. I. Krstic, R. Hansel, O. Romainczyk, J. W. Engels, V. Dötsch and T. F. Prisner, *Angew. Chem., Int. Ed. Engl.*, 2011, **50**, 5070–5074.
18. K. Bertrand, S. Reverdatto, D. S. Burz, R. Zitomer and A. Shekhtman, *J. Am. Chem. Soc.*, 2012, **134**, 12798–12806.
19. J. Hamatsu, D. O'Donovan, T. Tanaka, T. Shirai, Y. Hourai, T. Mikawa, T. Ikeya, M. Mishima, W. Boucher, B. O. Smith, E. D. Laue, M. Shirakawa and Y. Ito, *J. Am. Chem. Soc.*, 2013, **135**, 1688–1691.
20. K. Luby-Phelps, S. Mujumdar, R. Mujumdar, L. Ernst, W. Galbraith and A. Waggoner, *Biophys. J.*, 1993, **65**, 236–242.
21. S.-P. Williams, P. M. Haggie and K. M. Brindle, *Biophys. J.*, 1997, **72**, 490–498.
22. Z. Serber, W. Straub, L. Corsini, A. M. Nomura, N. Shimba, C. S. Craik, P. O. de Montellano and V. Dötsch, *J. Am. Chem. Soc.*, 2004, **126**, 7119–7125.
23. N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis and L. E. Kay, *J. Biomol. NMR*, 1999, **13**, 369–374.
24. Y. Ye, X. Liu, Z. Zhang, Q. Wu, B. Jiang, L. Jiang, X. Zhang, M. Liu, G. J. Pielak and C. Li, *Chemistry*, 2013, **19**, 12705–12710.
25. C. Li, G. F. Wang, Y. Wang, R. Creager-Allen, E. A. Lutz, H. Scronce, K. M. Slade, R. A. Ruf, R. A. Mehl and G. J. Pielak, *J. Am. Chem. Soc.*, 2010, **132**, 321–327.

26. E. A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb and S. A. Darst, *Cell*, 2001, **104**, 901–912.
27. M. Suzuki, J. Zhang, M. Liu, N. Woychik and M. Inouye, *Mol. Cell*, 2005, **18**, 253–261.
28. P. Selenko, Z. Serber, B. Gadea, J. Ruderman and G. Wagner, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 11904–11909.
29. P. Selenko and G. Wagner, *Nat. Methods*, 2006, **3**, 80–81.
30. K. Inomata, A. Ohno, H. Tochio, S. Isogai, T. Tenno, I. Nakase, T. Takeuchi, S. Futaki, Y. Ito, H. Hiroaki and M. Shirakawa, *Nature*, 2009, **458**, 106–109.
31. S. Ogino, S. Kubo, R. Umemoto, S. Huang, N. Nishida and I. Shimada, *J. Am. Chem. Soc.*, 2009, **131**, 10834–10835.
32. B. Bekei, H. M. Rose, M. Herzig, A. Dose, D. Schwarzer and P. Selenko, *Methods Mol. Biol.*, 2012, **895**, 43–54.
33. F. X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb and P. Selenko, *Nature*, 2016, **530**, 45–50.
34. C. Cedenio, H. Raveh-Hamit, A. Dinnyes and P. Tompa, *Adv. Exp. Med. Biol.*, 2015, **870**, 319–334.
35. R. Hansel, S. Foldynova-Trantirkova, F. Lohr, J. Buck, E. Bongartz, E. Bamberg, H. Schwalbe, V. Dotsch and L. Trantirek, *J. Am. Chem. Soc.*, 2009, **131**, 15761–15768.
36. R. Hansel, F. Lohr, L. Trantirek and V. Dotsch, *J. Am. Chem. Soc.*, 2013, **135**, 2816–2824.
37. D. Sakakibara, A. Sasaki, T. Ikeya, J. Hamatsu, T. Hanashima, M. Mishima, M. Yoshimasu, N. Hayashi, T. Mikawa, M. Walchli, B. O. Smith, M. Shirakawa, P. Guntert and Y. Ito, *Nature*, 2009, **458**, 102–105.
38. S. Reckel, J. J. Lopez, F. Lohr, C. Glaubitz and V. Dotsch, *ChemBioChem*, 2012, **13**, 534–537.
39. M. Renault, S. Pawsey, M. P. Bos, E. J. Koers, D. Nand, R. Tommassen-van Boxtel, M. Rosay, J. Tommassen, W. E. Maas and M. Baldus, *Angew. Chem., Int. Ed. Engl.*, 2012, **51**, 2998–3001.
40. M. Renault, R. Tommassen-van Boxtel, M. P. Bos, J. A. Post, J. Tommassen and M. Baldus, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 4863–4868.
41. S. Kubo, N. Nishida, Y. Udagawa, O. Takarada, S. Ogino and I. Shimada, *Angew. Chem., Int. Ed. Engl.*, 2013, **52**, 1208–1211.
42. N. G. Sharaf, C. O. Barnes, L. M. Charlton, G. B. Young and G. J. Pielak, *J. Magn. Reson.*, 2010, **202**, 140–146.

