Protein structure determination in living cells by in-cell NMR spectroscopy

Yutaka Ito[1,2,3], Daisuke Sakakibara[1,2], Atsuko Sasaki[1,2], Teppei Ikeya[1,4], Junpei Hamatsu[1,2], Tomomi Hanashima[1], Masaki Mishima[1,2], Tsutomu Mikawa[3], Markus Wälchli[5], Brian O. Smith[6], Masahiro Shirakawa[2,7] and Peter Güntert[1,4]

[1]Department of Chemistry, Tokyo Metropolitan University, Japan; [2]CREST/JST; [3]Research Group for Bio-supramolecular Structure-Function, RIKEN; [4]Institute of Biophysical Chemistry, J. W. Goethe-University Frankfurt, Germany; [5]Bruker BioSpin, Japan; [6]Division of Biochemistry and Molecular Biology, University of Glasgow, UK; [7]Department of Molecular Engineering, Kyoto University, Japan.

Introduction

In vivo observations of three-dimensional structures, structural changes, dynamics or interactions of proteins are essential for the explicit understanding of the structural basis of their functions inside cells. The non-invasive character and its ability to provide data at atomic resolution make NMR spectroscopy ideally suited for the task. Despite the interest in observing hetronuclear multi-dimensional NMR of macromolecules in living cells (in-cell NMR), it has not yet been confirmed whether the existing NMR methods for the structure determination of purified proteins can be applicable to proteins in living cells.

Results and Discussion

We use as a model system the *Thermus thermophilus* HB8 TTHA1718 gene product, a putative heavy metal binding protein consisting of 66 amino acids that was overexpressed in *E. coli* cells to a concentration of 3–4 mM. The stability of the live *E. coli* samples were examined under measurement conditions and confirmed that TTHA1718 in-cell NMR samples are stable for at least 6 hours (Figure 1a).

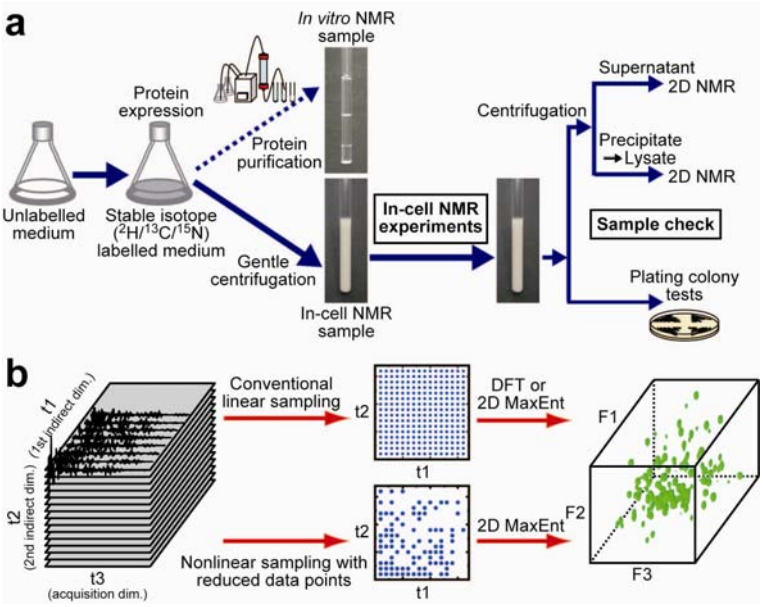

Figure 1. Experimental procedures for in-cell NMR experiments using *E. coli* cells. a, Scheme for checking the stability of *E. coli* cells under NMR measurement conditions. **b**, Rapid acquisition of 3D NMR spectra by employing a nonlinear sampling scheme. The 2D maximum entropy method was used for processing nonlinearly sampled dimensions.

First we performed backbone resonance assignment of TTHA-1718 in *E. coli* cells using six 3D triple-resonance NMR spectra. We applied a nonlinear sampling scheme for the indirectly acquired dimensions, which has been shown in combination with maximum entropy processing to provide significant time savings in the measurement of multi-dimensional NMR experiments (Figure 1b). All backbone resonances of the non-N-terminal and non-proline residues in TTHA1718 were assigned except for three residues in the hypothetical metal binding loop predicted from primary sequence analysis. Similarly we

assigned the majority of side-chain resonances by analysing three additional 3D triple-resonance NMR spectra.

Next, we performed NMR experiments of TTHA1718 in *E. coli* cells with selectively $^{1}H/^{13}C$-labelled methyl groups of Ala, Leu and Val, aiming at obtaining unambiguous long-range distance restraints involving methyl groups. Further NOE-derived distance restraints were obtained from 3D ^{15}N-separated NOESY and 3D ^{13}C-separated NOESY spectra measured on uniformly labelled *E.coli* samples.

The structure calculation of TTHA1718 in *E. coli* cells was performed with the program CYANA v3.0 on the basis of NOE-derived distance restraints, TALOS-derived backbone torsion angle restraints, and restraints for hydrogen bonds in α-helices and β-sheets. The resulting structure is well converged with a backbone RMSD of 0.96 Å to the mean coordinates (Figure 2a), and is very similar to the structure that was determined independently *in vitro* from a purified TTHA1718 sample (Figure 2b). The backbone RMSD between the in-cell and *in vitro* structures is 1.16 Å.

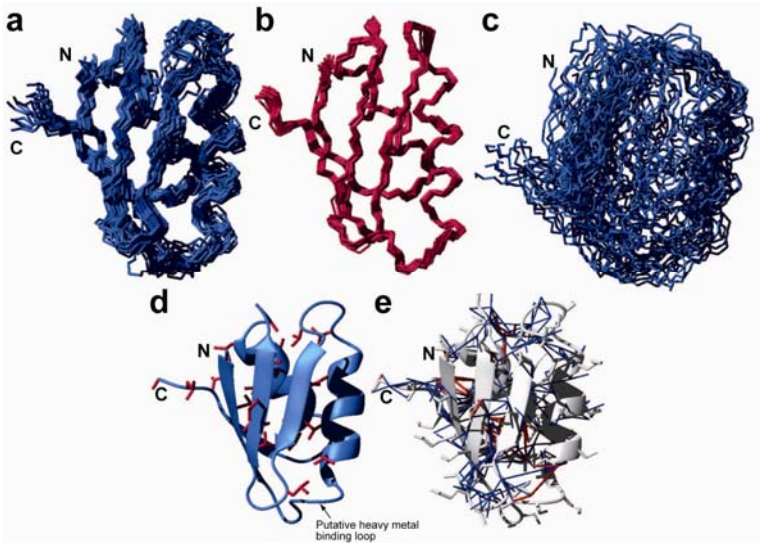

Figure 2. NMR solution structure of TTHA1718 in living *E. coli* cells. a, A superposition of the 20 final structures of TTHA1718 in living *E. coli* cells, showing the backbone (N, C^{α}, C') atoms. **b**, A superposition of the 20 final structures of purified TTHA1718 in vitro. **c**, A superposition of the 20 final structures of TTHA1718 in living *E. coli* cells calculated without distance restraints derived from NOEs involving methyl groups obtained in methyl-selectively protonated in-cell NMR samples. **d**, Secondary structure of TTHA1718 in living *E. coli* cells. The sidechains of Ala, Leu and Val residues whose methyl groups were labelled with $^{1}H/^{13}C$ are shown. **e**. Distance restraints derived from methyl group-correlated and other NOEs are represented in the ribbon model with thick and thin lines, respectively.

This result strongly suggests that high resolution 3D structures of proteins can be determined in solution in the cytoplasm of bacterial cells. Rapid data measurement using nonlinear sampling scheme with maximum entropy reconstruction, which provided spectra of much higher quality than conventional spectra recorded in the same short measurement time, and selective protonation at methyl groups (Figure 2d and 2e) to enable the observation of long-range unambiguous NOEs were key for the success of this study (Figure 2c).

By extending this approach it will become possible to study the conformations of proteins in detail and how they change in response to biological events in living environments. In particular, this approach provides the tools that will permit the effects of molecular crowding in the cytosol and the conformations of intrinsically disordered proteins to be investigated in living cells.