Characterization of proteins by in-cell NMR spectroscopy in cultured mammalian cells

Accepted Manuscript

Final version published online 19 May 2016
doi:10.1038/nprot.2016.061
https://www.nature.com/articles/nprot.2016.061

Characterization of proteins by in-cell NMR spectroscopy in cultured mammalian cells

Letizia Barbieri[1,2], Enrico Luchinat[1,3], Lucia Banci[1,2]

[1]Magnetic Resonance Center - CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy.
[2]Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy.
[3]Department of Biomedical, Clinical and Experimental Sciences, University of Florence, Viale Morgagni 50, 50134 Florence, Italy.
Correspondence should be addressed to L.Banci (banci@cerm.unifi.it).

In-cell NMR spectroscopy is a unique tool to characterize biological macromolecules in their physiological environment at atomic resolution. Recent progresses of NMR instruments and sample preparation methods allow functional processes, such as metal uptake, disulfide bond formation, protein folding, to be analyzed by NMR in living, cultured human cells. This protocol describes the necessary steps to overexpress one or more proteins of interest inside human HEK293T cells and explains how to setup in-cell NMR experiments. The cDNA is transiently transfected as a complex with a cationic polymer, and protein expression is carried on for 2-3 days, followed by the NMR sample preparation. ^{1}H and ^{1}H-^{15}N SOFAST-HMQC NMR experiments can be acquired in less than 2 hours, ensuring cell viability. Uniform ^{15}N labeling and amino acid specific (e.g. cysteine, methionine) labeling schemes are possible. The entire procedure takes 4 days from cell culture seeding to NMR data collection.

INTRODUCTION

In-cell NMR spectroscopy allows the study of biomolecules, in particular proteins, inside living cells. It is therefore a powerful method for acquiring structural and functional information of biomolecules at atomic resolution in their natural, physiological environment[1–5]. In-cell NMR has been successfully employed on bacterial cells to analyze protein structure[6], dynamics[7,8] and interactions[9–11]. Although prokaryotic cells have some advantages, like being easy to handle, growing fast, and expressing exogenous proteins at high levels, to be meaningful eukaryotic proteins should be characterized in eukaryotic cells. Indeed, prokaryotic organisms exhibit a limited range of biological activities and many of the cellular processes that define the challenging questions in biological research are absent in bacteria. Post-translational protein modifications regulate biological activities in eukaryotes, but are much less common in prokaryotic organisms. The presence of organelles and the requirement of regulated cellular transport constitutes another characteristic of eukaryotic cells. Compartmentalization causes the creation of sub-cellular environments with different physical and biological properties and the effects that these compartments exert on the protein conformation, state and/or function need to be properly addressed.

To date, several approaches have been used to have the proteins of interest properly labeled and located inside the cytoplasm of eukaryotic cells. Microinjection, cell-penetrating peptides, pore-forming toxin, and electroporation have been the most commonly used methods[3,10,12–14]. The large size of *Xenopus laevis* oocytes allows microinjection of isotopically labeled proteins, produced in bacteria, into the cytosol[12]. In this system, proteins phosphorylation and structures of nucleic acids have been investigated[15–18]. Alternatively, the isotopically labeled proteins could be delivered into the eukaryotic cytoplasm via CPP (cell-penetrating peptides) covalently linked to the protein. The proteins are subsequently released from CPP by endogenous enzymatic action or by autonomous reductive cleavage[3]. Ogino *et al.* exploited the pore-forming toxin streptolysin O to import the protein thymosin β4 in human cells[13]. One advantage of this approach, in contrast to the CPP strategy, is that no modifications to the protein are required for intracellular sample delivery[19]. Recently, protein electroporation has been developed; this approach involves applying an electric

pulse to cells so to insert the purified protein of interest into the cytosol[10,14].

All these methods rely on heterologous protein expression and purification, and its subsequent insertion in the cell. This implies that several aspects cannot be easily addressed, such as the steps of protein maturation including folding, cofactor binding, cysteine oxidation etc., as the proteins expressed and purified from bacterial cells are already folded/oxidized and usually need additional treatment prior to insertion, making harder to obtain meaningful information from in-cell NMR data. Therefore, a strategy for producing, labeling and studying proteins directly in eukaryotic cells is desirable. Intracellular expression strategies for in-cell NMR have been successfully applied to yeast and insect cells[20,21]. Such approaches would be especially suited for proteins that 1) have tendency to aggregate *in vitro* at high concentrations; 2) undergo co-translational processes for correct folding or intracellular targeting; 3) are sensitive to the redox properties of the environment (e.g. require thiol-based reducing agents and/or absence of oxygen). Furthermore, human proteins should be investigated in a human cellular environment, to closely match the physiological conditions in terms of molecular composition.

Here, we describe the step-by-step procedure for the preparation of samples of Human Embryonic Kidney 293T (HEK293T) cells expressing isotopically labeled proteins, and for the recording of in-cell NMR experiments. This protocol relies on transiently transfecting the cells with one or more genes of interest, using a vector designed for high level constitutive cytoplasmic protein expression. Isotope labeling is performed by providing a growth medium isotopically labeled at the time of transfection. Co-transfection of two or more genes can be performed, resulting in the simultaneous expression and labeling of the proteins. After protein expression, in-cell NMR data is collected to obtain information on the intracellular conformation of the protein(s). Compared to the protein insertion strategies, intracellular expression has the advantage of providing higher physiological relevance especially when applied to study processes highly dependent on the chemical environment, whereas a disadvantage lies in the non-specific isotopic enrichment during protein expression, which causes cellular background signals to be detected together with the protein of interest, requiring further data processing.

Despite the general limitation of in-cell solution NMR, which cannot detect globular proteins if they interact with large cellular components (see Troubleshooting of Step 33), this protocol should be applicable to a broad range of soluble cytoplasmic proteins, and can be adapted to proteins localized in different cellular compartments.

Applications

The protocol for protein overexpression in human HEK293T cells for in-cell NMR studies has been successfully applied to a number of systems[5,22–27]. Several functional processes, including metal uptake[5,24], disulfide bond formation[5,22,24,26], protein folding[22,26], have been characterized for different proteins with different features and behaviours, and in different cellular compartments[23]. For superoxide dismutase 1 (SOD1) and for a number of its familial Amyotrophic Lateral Sclerosis (fALS)-linked mutants, the post-translational modifications occurring after the protein synthesis were studied by analyzing the protein state as a function of the growth conditions[5,24]. The maturation process of SOD1 involves zinc binding, homodimer formation, copper uptake and oxidation of the intrasubunit disulfide bond. By treating the cells differently during protein expression, the various protein conformations were observed, from the immature, metal-free form to the various metallated and/or oxidized forms to the final mature protein[5].

In cells grown without supplements of metal ions, essentially only the monomeric, metal-free, reduced SOD1 (apo-SOD1) is present. Addition of Zn^{2+} to the culture medium eliminated the apo-SOD1 species and only the dimeric species with one Zn^{2+} ion bound to each subunit (E,Zn-SOD1) is formed. Therefore, Zn^{2+} ions are efficiently taken up by cells in culture and bind specifically to the native binding site of dimeric SOD1 in stoichiometric amounts. Copper incorporation of SOD1 in eukaryotes is known to be dependent on the CCS protein. When both proteins are overexpressed, the SOD1 intrasubunit disulfide bond is partially oxidized. When cells, after overexpression of both SOD1 and CCS, are incubated with Cu(II), almost all SOD1 is in the Cu(I),Zn-bound form, at variance with cell samples with basal CCS level. Moreover, the intrasubunit disulfide bridge of SOD1 is completely oxidized. This behaviour indicates that *in vivo* CCS has a role in promoting both copper incorporation in SOD1 and disulfide bond formation[28].

SOD1 maturation pathway has also been investigated for a series of fALS-linked SOD1 mutants revealing that, for a subset of mutants, impaired zinc binding triggers the accumulation of an unstructured species in the cytoplasm[24]. Co-expression of CCS prevented SOD1 misfolding, and restored the correct maturation pathway. This application highlights the advantage of intracellular protein expression over the protein delivery approaches. In the apo state, these

SOD1 mutants are intrinsically destabilized, are prone to misfolding and aggregation *in vitro*, and cannot reach the high concentration needed for efficient intracellular delivery. Therefore, the approach described here is especially valuable for studying protein misfolding events which are difficult to control *in vitro*.

We have also characterized the folding and redox properties of Mia40, an oxidoreductase involved in the mitochondrial protein import process[22]. The in-cell NMR spectra revealed that Mia40 in the cytoplasm is present mostly in a folded state, which is not competent for mitochondrial import. The dependence of Mia40 redox and folding state on the overexpression of Glutaredoxin1 (Grx1) and Thioredoxin1 (Trx1) has also been investigated[22]. The amount of intracellular folded Mia40 decreased when either Grx1 or Trx1 were co-expressed, showing that both thiol-disulfide regulating proteins affect the Mia40 redox state. Grx1 keeps Mia40 reduced more effectively than Trx1, suggesting that the two proteins differentially regulate the mitochondrial import of Mia40.

Experimental design

While this protocol was optimized on HEK293T cells, in principle protein expression at sufficient levels can be obtained in other mammalian cell lines. However, some steps will likely have to be modified. Specifically, the transfection method and the choice of the promoter for constitutive gene expression should be optimized for each cell line. As a rule of thumb, cells that have been used successfully as protein expression systems (e.g. HEK293T[29] cells, Chinese Hamster Ovary (CHO)[30] cells and simian-derived COS cells[31]) should be able to reach high transfection efficiency and sufficient protein levels for in-cell NMR experiments.

The genes of interest need to be cloned in a vector suitable for high level constitutive protein expression in mammalian cells. In this protocol, the pHLsec plasmid is used[29,32]. The gene of interest is transcribed by a CAG promoter, which is very active in HEK293T cells and allows high protein levels to be reached. The pHLsec vector is designed for expressing secreted proteins, however the secretion sequence can be removed during the cloning, allowing the native proteins to be expressed in the cytoplasm[5]. In order to ensure successful transfection and protein expression, the transfected DNA has to be endotoxin-free, as bacterial endotoxins induce inflammatory reactions of the mammalian immune system. Usually, endotoxin-free Midi and Maxi DNA preparations are preferred, as they provide enough DNA for ~15 and ~30 in-cell NMR samples, respectively. The DNA is transfected as a complex with the cationic polymer polyethylenimine (PEI), which ensures high transfection efficiency and high copy number per cell[29].

For each protein, the optimal expression conditions (in terms of time and amount of DNA) may have to be screened, either by Coomassie-stained SDS-PAGE or by Western Blotting of the cell lysates. Usually, the highest expression levels are reached between 48 and 72 hours after transfection. It is possible to co-express two proteins simultaneously by mixing the two DNAs in the transfection procedure. Optimal co-expression conditions have to be screened by varying the amount of DNA of each construct. If desired, the protein expression level can be decreased by diluting the DNA encoding the protein of interest with an empty DNA vector (the total amount of DNA is kept constant). The screening can be done on scaled-down cultures (e.g. in T25 flasks or 6-well dishes). To quantify the expression levels, the lysates obtained for each condition are run on a reducing SDS-PAGE together with serial dilutions of a purified protein sample at known concentration, which provides a calibration curve. Each sample should be loaded at different dilutions for better quantification. Protein detection can be performed by densitometry analysis on the Coomassie-stained gel. Alternatively, especially for lower expression levels, the proteins can be detected by Western Blotting and densitometry analysis.

Some applications may require additional reagents to be added to the cell culture at some point of the sample preparation. As an example, some proteins need cofactors to fold properly and/or to reach the functional state. Such cofactors can be either small molecules or metal ions. In order to avoid cellular toxicity or decreased protein expression, the optimal concentration of cofactor(s) and timing of the addition(s) should be tested. For example, copper ions are toxic at high concentration and inhibit protein expression, so they have to be added at a later stage following protein expression. Other transition metals, such as zinc, do not seem to affect protein expression and can be added at the time of transfection. Protein expression and cell viability can be checked by Western Blot and Trypan Blue staining, respectively, while the effects on the protein folding/functional state can be observed by NMR experiments on intact cells or lysates.

A prerequisite to perform heteronuclear NMR experiments is that NMR-active isotopes (^{13}C and ^{15}N) must be incorporated into the overexpressed proteins. For this purpose, the growth medium needs to be replaced at the time of transfection with a medium isotopically labeled. Labeled media suitable for mammalian cell growth are commercially available (Bioexpress 6000, CIL), and cost around 5,500 USD/l

(U-^{15}N) and 11,000 USD/l (U-^{13}C,^{15}N). Despite the high initial costs, following our protocol only 20 ml of labeled medium are required to obtain one in-cell NMR sample, resulting in 110 USD/sample (^{15}N) and 220 USD/sample (^{13}C,^{15}N). In most cases, the labeled media can be recovered after protein expression, centrifuged to remove cell debris and residual DNA:PEI complex, and used again to prepare a second cell sample, further reducing the cost per sample. For some applications, typically when some protein signals fall in a spectral region free from cellular background (for example, histidine ring N-^{1}H protons[5,24]; isolated ^{1}H signals arising from the protein hydrophobic core[22,27]), unlabeled samples can be prepared for ^{1}H-only in-cell NMR experiments, resulting in the cost per sample being greatly reduced. Amino acid-selective labeling schemes can also be used, by preparing a home-made cell growth medium following the composition of the commercial unlabeled medium and adding the isotopically labeled amino acid(s) according to the desired labeling scheme. During protein expression, the amount of serum is lowered to 2%, to prevent excessive isotopic dilution and to minimize non-specific isotopic incorporation, by decreasing the cell growth rate[32,33].

In our protocol, 3 mm Shigemi NMR tubes are used to minimize the amount of cells needed and cut the sample preparation costs. On average, an NMR sample obtained from one confluent T75 flask (75 cm^2) contains $3.5\times10^7 \pm 1.1\times10^7$ cells, which occupy 130 ± 30 µl when spun down; this is enough to fill the tube section corresponding to the active volume of the NMR instrument. The resulting pellet contains ~50% (v/v) cells (average cell volume of 2 pl). 3 mm Shigemi tubes can be used in 5 mm probes with a proper spinner/adapter. If more cells are available, larger tubes (5 mm normal/Shigemi tubes) may be used to increase the signal to noise ratio (S/N) of the experiments, as shown elsewhere[34,35].

After the cells are detached from the culture flask, the in-cell NMR experiments should be performed as quickly as possible to minimize cell death. In order to overcome the intrinsic low sensitivity of NMR, some pulse sequences have been optimally designed to maximize the S/N ratio per unit time. For example, ^{1}H-^{15}N SOFAST-HMQC experiments are preferred instead of the classical HSQC variants. Usually, more than 90% of HEK293T cells are still viable after 2-2.5 hours at 37°C, which is longer than needed to collect one SOFAST-HMQC spectrum (1 h) plus several 1D ^{1}H NMR spectra. Longer experimental times result in cell death due to O_2 and/or nutrients depletion, and build-up of acidic metabolic products. For these reasons, NMR experiments should be planned and set up in advance whenever possible.

MATERIALS

REAGENTS

- HEK293T cell line (ATCC CRL-3216) **!CAUTION** the cell lines used should be regularly checked to ensure they are authentic and are not infected with mycoplasma.
- pHLsec plasmid[29] containing the gene of interest; empty pHLsec plasmid.
- Phosphate-Buffered Saline (PBS, Life Technologies, cat. no. 10010)
- Dulbecco's Modified Eagle Medium, high glucose (DMEM, Life Technologies, cat. no.10313-021)
- Fetal Bovine Serum (FBS, Life Technologies, cat. no. 10270)
- L-Glutamine (Life Technologies, cat. no. 25030)
- Penicillin-Streptomycin (10,000 U/ml) (Life Technologies, cat. no. 15140-122)
- Trypsin-EDTA (0.05%) (Life Technologies, cat. no. 25300-054)
- MEM Non-Essential Amino Acids Solution (100X) (NEAA, Life Technologies, cat. no. 11140-035)
- Polyethylenimine (PEI), branched, avg. M_w 25 kDa (Sigma-Aldrich, cat. no. 408727)
- BioExpress 6000 (U-^{15}N, 98%) (CIL, cat. no. CGM-6000-N)
- HEPES (Sigma-Aldrich, cat. no. H4034)
- D-(+)-Glucose (Sigma-Aldrich, cat. no. G8270)
- D_2O (Sigma-Aldrich, cat. no. 453366)
- Ethylenediaminetetraacetic acid disodium salt solution (EDTA, Sigma-Aldrich, cat. no. E7889)
- 4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF Hydrochloride protease inhibitor, PanReac AppliChem, cat. no. A1421,0001)
- Trypan Blue Solution (0.4%) (Sigma-Aldrich, cat. no. T8154) **!CAUTION** Hazard statement(s): H350, may cause cancer.

Optional reagents:

- L-Methionine (Methyl-13C, 99%) (CIL, cat. no. CLM-206-PK)
- L-Cysteine (U-13C3, 97-99%; 15N, 97-99%) (CIL, cat. no. CNLM-3871-0.1)
- $CaCl_2{\cdot}2H_2O$ (Merck Millipore, cat. no. 1.02382)
- $Fe(NO_3)_3{\cdot}9H_2O$ (Sigma-Aldrich, cat. no. 216828)
- $MgSO_4{\cdot}7H_2O$ (Sigma-Aldrich, cat. no. M2643)
- KCl (Merck Millipore, cat. no. 1.04936)
- $NaHCO_3$ (Carlo Erba, cat. no. 478537)

- NaCl (Sigma-Aldrich, cat. no. 31434)
- NaH_2PO_4 (Merck Millipore, cat. no. 1.06346)
- L-Arginine·HCl (Sigma-Aldrich, cat. no. A5006)
- L-Cysteine·2HCl (Sigma-Aldrich, cat. no. C7352)
- Glycine (Sigma-Aldrich, cat. no. G7126)
- L-Histidine·HCl·H_2O (PanReac AppliChem, cat. no. A1591)
- L-Isoleucine (Sigma-Aldrich, cat. no. I7403)
- L-Leucine (Sigma-Aldrich, cat. no. L8000)
- L-Lysine·HCl (Sigma-Aldrich, cat. no. L5626)
- L-Methionine (Sigma-Aldrich, cat. no. 64320)
- L-Phenylalanine (Sigma-Aldrich, cat. no. P5482)
- L-Serine (Sigma-Aldrich, cat. no. 84959)
- L-Threonine (Sigma-Aldrich, cat. no. T8625)
- L-Tryptophane (Sigma-Aldrich, cat. no. 93660)
- L-Tyrosine·2Na·2H_2O (Sigma-Aldrich, cat. no. 93830)
- L-Valine (Sigma-Aldrich, cat. no. 94620)
- Choline Chloride (Sigma-Aldrich, cat. no. C7527)
- Folic Acid (Sigma-Aldrich, cat. no. F7876)
- D-Pantothenic Acid·½Ca (Sigma-Aldrich, cat. no. 25,972-1)
- Pyridoxine·HCl (Sigma-Aldrich, cat. no. P5669)
- Niacinamide (Sigma-Aldrich, cat. no. N-3376)
- *myo*-Inositol (Sigma-Aldrich, cat. no. I7508)
- Riboflavin (Sigma-Aldrich, cat. no. R-4500)
- Thiamine·HCl (Merck Millipore, cat. no. 1.08181)
- Purified proteins, either unlabeled or labeled according to the in-cell NMR experiments (i.e. U-^{15}N or amino acid-specific)

EQUIPMENT

- All the equipment necessary for maintaining cultured mammalian cells.
- All the equipment to perform SDS-PAGE and Western-Blotting procedures.
- Access to a high-field NMR spectrometer equipped with a cryogenically cooled probe-head is a requirement for successful in-cell NMR analysis. The data shown here were obtained at a 950 MHz Bruker Avance spectrometer equipped with a 5 mm TCI CryoProbe.
- NMR data-processing software (e.g. TOPSPIN for Bruker machines).
- Visualization and analysis software for NMR spectra (e.g. TOPSPIN, CARA, SPARKY and others).
- Cell culture flasks, T75 (Greiner Bio-one, cat. no. 658175).
- Burker chamber (Sigma-Aldrich, cat. no. BR718920).
- 3 mm Shigemi NMR tubes (Cortecnet, cat. no. BMS-003).
- Liquid N_2.

REAGENT SETUP

Complete DMEM Add 1x L-Glutamine, 1x Penicillin-Streptomycin and 10% FBS to DMEM. Can be stored at 4°C for 1 month.

Unlabeled expression medium Add 1x L-Glutamine, 1x Penicillin-Streptomycin, 1x NEAA and 2% FBS to DMEM. Can be stored at 4°C for 1 month.

Labeled expression medium Add 1x Penicillin-Streptomycin and 2% FBS to BioExpress 6000. Freshly prepare the amount needed for each experiment.

"Home-made" DMEM Weigh every component following the recipe of the commercial DMEM (e.g. Sigma-Aldrich, Life Technologies) and dissolve them in H_2O. Adjust the pH of the solution to 7.5. Sterilize the medium by filtration (0.22 µm). The medium can be stored at -20°C for 1 year. For the expression, add 1x Penicillin-Streptomycin, 2% FBS to DMEM and the labeled amino acid following the recipe. Add 1x L-Glutamine and 1x NEAA only if they do not contain the amino acid chosen for labeling. Freshly prepare the amount needed for each experiment.

PROCEDURE

Splitting of confluent cells (day 1) ● TIMING 30 min

1| Remove the medium from a confluent T75 flask of HEK293T cells (containing ~1-3 x 10^7 cells). One T75 flask is sufficient to inoculate 4-10 flasks for in-cell NMR experiments (see Step 7).

2| Wash the cells twice with 7 ml of PBS.

3| Add 2 ml of trypsin/EDTA, incubate 2 min at room temperature to detach the cells.

4| Inactivate trypsin with 20 ml of complete DMEM.

5| Centrifuge cells at 800 g for 5 minutes at room temperature, discard the supernatant.

6| Resuspend the cell pellet in 10 ml of complete DMEM, carefully breaking up any clumps of cells.

7| If transfecting the cells on day 2, for each in-cell NMR sample inoculate 2.2 ml of cell suspension in a new T75 flask, and add 20 ml of complete DMEM. Alternatively, the DNA transfection can be performed on day 3. If that is the case, inoculate 1 ml of cell suspension per in-cell NMR sample. See **Figure 1** for the alternative timing.

8| Place the inoculated flasks at 37°C in a humidified atmosphere containing 5% CO_2.

Transient transfection with the gene(s) of interest / with the empty vector (day 2) ● TIMING 30 min

9| To 2.5 ml of medium (depending on the chosen labeling scheme, use either DMEM, ^{15}N-labeled medium or home-made medium with specific labeled amino acids) without serum add 25 µg of endotoxin-free DNA.

10| To 2.5 ml of medium without serum (the same as in Step 9) add 50 µg of PEI.

11| Slowly add the PEI solution to the DNA solution.

▲CRITICAL STEP To ensure proper transfection efficiency, add PEI to DNA, not *vice versa*. Do not shake the tube.

12| Wait for 20 minutes in order to allow the formation of the DNA-PEI complex.

13| Discard the old medium from the T75 flask and carefully eliminate any trace of medium

14| Add the DNA-PEI mixture to the cells and subsequently add 15 ml of the same medium as in Step 9 supplemented with 2% FBS and antibiotics. Non-toxic cofactors / metal ions can be added to the expression medium during this step. If toxic compounds / metal ions are required, they should be added after protein expression, prior to cell collection.

15| Place the flask at 37°C in a humidified atmosphere containing 5% CO_2. Carry on protein expression for 2 or 3 days (see **Figure 1**).

Setup of the NMR experiments (day 4) ● TIMING 1 h

16| Set the experimental temperature at the NMR spectrometer. Typically, the temperature is set to 310 K, which is close to the physiological temperature, and provides 2-3 hours of sample stability. Lower temperatures may increase the sample stability, but they usually cause noticeable line broadening, due to the increased viscosity of the environment.

17| Prepare an empty dataset at the NMR spectrometer and set the parameters for each NMR experiment by following the options A and B. This step should be performed at any time prior to the preparation of the in-cell NMR sample, in order to minimize the delay between cell collection and data acquisition.

There are several variants of NMR experiments which are suitable for short-lived samples such as living HEK293T cells put in an NMR tube. Here we describe those most commonly used in the reported applications of this protocol. Typically, samples are first analyzed with a 1D ^{1}H NMR experiment with excitation sculpting for water suppression (Bruker zgesgp pulse program)[36](option A). In the case that a sizable number of ^{1}H resonances of the protein of interest fall in regions of the ^{1}H spectrum free of signals from the cellular background (i.e. >10 ppm or <0.5 ppm), analysis by ^{1}H NMR on simply unlabeled cell samples can be performed. In such cases, a 1D ^{1}H NMR WATERGATE experiment with a 3-9-19 binomial pulse train for water suppression provides good sensitivity (Bruker p3919gp pulse program)[37].

With ^{15}N labeled samples, after the acquisition of a 1D ^{1}H NMR spectrum, a ^{1}H-^{15}N SOFAST-HMQC NMR experiment (Bruker sfhmqcf3gpph pulse program)[38] is acquired (option B), which will be the major source of structural information. The SOFAST-HMQC sequence relies on band-selective shaped pulses for ^{1}H excitation and refocusing. Both the pulse length and the offset frequency can be adjusted depending on the desired ^{1}H excitation region. In our standard setup, the pulse lengths and offset are calculated for the amide region of folded proteins. The inter-scan delay can be optimized, within the safe operating limits of the instrument, by measuring the obtained S/N per unit time at different delays. Optionally, non-uniform sampling schemes can be used, either to reduce the acquisition time for each experiment (provided the S/N is high enough) or to improve the resolution along the ^{15}N dimension (within the limits of ^{15}N transverse relaxation)[3].

(A) Parameters for 1D ^{1}H NMR experiments

(i) Center the ^{1}H carrier frequency on the water signal at ~4.7 ppm

(ii) For the zgesgp pulse program, set the spectral width to 20 ppm and a 1000 µs 180° square pulse for water suppression. Set an inter-scan delay of ≥ 1 second. Acquire the spectrum with 128 scans.

(iii) For the p3919gp pulse program, set the spectral width to ≥ 25 ppm, and adjust the delay for binomial water suppression so that the maximum excitation is centred at the chemical shifts of the signals of interest. Set an inter-scan delay of ≥ 1 second. Acquire with ≥ 512 scans.

(B) Parameters for 2D ^{1}H-^{15}N SOFAST-HMQC experiments

(i) Center the 1H carrier frequency on the water signal at ~4.7 ppm and the 15N frequency at ~118 ppm.

(ii) Set the spectral width to 15 ppm for 1H and 50 ppm for 15N, 2048 points and 128 increments (acquisition times of 67.4 ms and 13.3 ms at 950 MHz, respectively).

(iii) Center the band-selective 1H shaped pulses (use a PC9 shape for excitation and a Reburp shape for refocusing) at 8.5 ppm or higher, to avoid excitation of the water resonance.

(iv) Calculate the power level for the PC9 shaped pulse (either 90° or 120° Ernst angle can be used). Set the duration to 1620 µs (at 950 MHz).

(v) Calculate the power level for the Reburp shaped pulse, set the duration to 896 µs (at 950 MHz). At different magnetic fields, both pulse lengths need to be adjusted with the following formula: pnew = pold × (Bold/Bnew).

(vi) Set an inter-scan delay of 300 ms. Acquire with 64 scans per increment.

Preparation of the NMR sample ● TIMING 1 h

18| Remove the old medium from the T75 flask.

? TROUBLESHOOTING

19| Wash the cells twice with 7 ml (each) of PBS.

20| Detach the cells with 2 ml of trypsin/EDTA, incubate 5 min at room temperature to detach the cells.

▲CRITICAL STEP trypsin may take slightly longer to detach transfected cells. If necessary, incubate the cells at 37°C.

21| Inactivate trypsin with 20 ml of complete DMEM.

22| Centrifuge the cells at 800 g for 5 minutes at room temperature, discard the supernatant.

23| Wash cells with 10 ml of PBS to remove the residual medium.

24| Centrifuge the cells at 800 g for 5 minutes at room temperature, discard the supernatant.

25| Transfer the cells to a 1.5 ml capped tube.

26| Resuspend the cell pellet in 180 µl of NMR buffer (DMEM supplemented with 70 mM HEPES, 90 mM Glucose, 20% (v/v) D_2O). This formulation was optimized for increased buffering capacity and for compensating for the faster glucose consumption, due to the high cell density in the NMR tube.

▲CRITICAL STEP DMEM is buffered by CO_2/HCO_3^-. To avoid the formation of CO_2 bubbles in the NMR tube, after HEPES addition, remove the CO_2 by bubbling air for 10 minutes.

27| Take 10 µl of the cell suspension for measuring the cell viability before the NMR experiments by Trypan Blue staining (Steps 45-48). The aliquot can be stored on ice for 1-2 hours.

28| With a long glass pipette, transfer the cell suspension to the bottom of a 3 mm Shigemi NMR tube; do not use the inner plunger, close the tube with a standard 3 mm cap.

29| Allow the cells to settle at the bottom of the tube. This step can be made quicker by slowly spinning the tube with a hand-operated centrifuge for ~3 minutes. Do not remove the supernatant.

Acquisition of the NMR spectra ● TIMING 1-3 h

30| Load a data set with the desired NMR experiments (see Step 17)

31| Insert the in-cell NMR sample into the magnet, check the tuning of the probe head and shim the magnet.

▲CRITICAL STEP In our experience, the automated shimming procedure of modern Bruker instruments (the 'topshim' routine) applied on the cell samples yields results as good as on the same Shigemi tube filled with aqueous buffer and the

inner plunger. In practice, any residual field inhomogeneity seems not to be critical compared to the signal broadening caused by the protein of interest being inside the cells. Water suppression could pose some issues on poorly shimmed samples. However, SOFAST-HMQC spectra can be optimized for such samples by moving the ^{1}H excitation window (i.e. the offset frequency of the band-selective ^{1}H shaped pulses) further away from the water signal.

32| Calibrate the ^{1}H 90° hard pulse length.

33| Adjust the ^{1}H power levels in each pulse sequence according to the ^{1}H hard pulse.

34| Acquire the NMR spectra. Process the data as described in Box 1. A quick processing can be performed to check the outcome of the experiment, while the background subtraction can be performed at a later stage.
? TROUBLESHOOTING

Recovery of the cell sample ● TIMING 10 min

35| After the NMR experiment, resuspend the cells in the supernatant inside the NMR tube with a glass pipette and transfer the suspension in a 1.5 ml capped tube.

36| Take 10 µl of the cell suspension for measuring the cell viability after the NMR experiments by Trypan Blue staining (Steps 45-48). The aliquot can be stored on ice for 1-2 hours.

37| Centrifuge the cells at 800 g for 5 minutes at room temperature.

38| Collect the supernatant, to be analyzed with NMR for protein leakage by performing Steps 30-35.
? TROUBLESHOOTING
■ **PAUSE POINT** the cell pellet and the supernatant can be stored at –20°C before preparation of the cell lysate.

Preparation of the cell lysate ● TIMING 1.5 h

39| Resuspend the cell pellet in 150 µl of PBS buffer + 0.5 mM EDTA + AEBSF.

40| Lyse the cells by applying 8-10 freeze–thaw cycles alternating between liquid N_2 and warm water.

41| Centrifuge the lysate at 16000 g for 1 hour, 4 °C and collect the supernatant.

42| Analyse the cleared cell lysate by NMR by performing Steps 30-35. A conventional 3 mm NMR tube can be used.
? TROUBLESHOOTING

43| (optional) Perform further analysis on lysate or on supernatant, e.g. by SDS-PAGE or Western Blot.

Trypan Blue test ● TIMING 10 min

▲**CRITICAL** Perform the Trypan Blue test on the cell aliquots taken before and after the NMR experiments (Steps 27 and 37):

44| Dilute the 10 µl cellular suspension in 500 µL of DMEM without serum.

45| Add 100 µl of Trypan Blue Stain.

46| Wait for 5 minutes.
▲**CRITICAL STEP** Do not exceed 5 minutes, otherwise every cell becomes stained.

47| Put 10 µl in a Burker chamber and count the number of blue (B) and white (W) cells. Calculate the cell viability as W/(B+W). Ideally, cell viability should remain higher than 90% after the NMR experiments.

***In vitro* NMR data collection**

48| (optional) Acquire the same set of NMR experiments on purified protein samples by performing Steps 30-35, in order to compare the in-cell and *in vitro* NMR data.

Box 1 | Processing of the 2D NMR experiments

Typically, 2D ^{1}H-^{15}N NMR spectra are processed by multiplying both dimensions by a square sine bell function (90° shift), zero-filled and Fourier transformed in both dimensions. In some cases, amide resonances arising from the in-cell NMR sample will relax faster than the acquisition time (in one or both spectral dimensions). In such cases, the spectra can be improved by decreasing the number of effective data points used in the Fourier transform, prior to zero-filling. As this choice can be made during the processing step, it is suggested to always acquire the NMR spectra with the highest possible number of points in both dimensions (compatibly with the experimental time and with the safety limits for decoupling during acquisition), in order to maximize the spectral resolution for slow-relaxing signals. If needed, some points can then be discarded later.

In the case of isotopically labeled cell samples, during protein expression the NMR-active isotope(s) will also be incorporated into all the other molecules of the cells. Therefore, ^{15}N-labeled cell samples give rise to background amide signals in the 2D NMR spectra, originating mostly from other proteins and abundant peptides, such as glutathione. These background signals can be greatly reduced in the 2D NMR spectra by subtracting two Fourier-transformed spectra, one acquired on the in-cell NMR sample containing the protein(s) of interest, the other acquired on a control in-cell NMR sample, prepared exactly in the same conditions as the previous but transfecting the cells with the empty DNA vector (**Fig. 2**). The two experiments need to be set up and processed using identical parameters (except for the power levels of the shaped pulses, which should be adjusted according to the ^{1}H hard pulse length). Therefore, it is advisable to plan the set of 2D NMR experiments always combined with a control in-cell NMR spectrum. All these spectra should be collected with identical experimental parameters such as temperature, spectral width, number of increments and shape pulse profiles. Thus, a single control in-cell NMR sample can be used for processing different in-cell NMR experiments, provided that the cell culture conditions and experimental set up have not changed along the collection of the various experiments.

? TROUBLESHOOTING

Step 18:
Problem: Most cells are detached from the flask and are floating in the medium.
Possible solutions: Cytotoxic effects may arise from added cofactors. Determine the optimal conditions by lowering the cofactor concentration.
The expressed protein may be toxic. Tune the amount of transfected DNA by diluting the plasmid containing the gene with an empty vector.

Step 35:
Problem: No signals of the protein of interest are observed in the in-cell NMR spectra.
Possible solution: Soluble, globular proteins are sometimes not detectable by in-cell solution NMR due to interactions with other cellular components, which slow down the average tumbling rate. For example the human proteins Grx1, Trx1 and profilin 1 do not give rise to detectable signals in the NMR spectra (**Supplementary Fig. 1**)[22,25]. To determine whether this is the case, lyse the cells and acquire the same NMR spectra on the cleared cell lysate. If the protein is clearly detected in the cell lysate, but not in the cells, a likely reason is the occurrence of transient interactions with large intracellular molecules. In such cases, a mutational strategy may be followed, to selectively disrupt transient interactions. This approach has proven to be successful in making human profilin 1 detectable by in-cell solution NMR[25].

Step 39:
Problem: The protein NMR signals are detected in the supernatant: the protein has leaked from the cells. This is likely a consequence of high cell mortality, which can be assessed with the Trypan Blue test.
Possible solutions: Try to be as quick as possible in the preparation of the NMR sample by reducing the dead-times in the procedure.
Reduce the acquisition time of the NMR experiments.
The over-expressed protein might be produced at levels that are toxic to the cells. Tune the amount of transfected DNA by diluting the plasmid containing the gene with an empty vector; this will decrease the number of copies of the plasmid entering each cell.

Step 43:
Problem: No signals of the protein of interest are observed in the NMR spectra of the cleared cell lysate.
Possible solution: the protein of interest is not expressed to sufficient levels. Check the presence of the protein with SDS-PAGE or Western Blotting. Vary the amount of DNA in the transfection procedure.

● TIMING
Steps 1–8, splitting of confluent cells: 30 min
Steps 9–15, transient transfection with the gene(s) of interest or empty vector: 30 min
Steps 16 and 17, setup of the NMR experiments: 1 h
Steps 18–29, preparation of the NMR sample: 1 h
Steps 30–34, acquisition of the NMR spectra: 1–3 h
Steps 35–38, recovery of the cell sample: 10 min
Steps 39–43, preparation of the cell lysate: 1.5 h
Steps 44–48, Trypan Blue test: 10 min
Box 1, processing of the 2D NMR experiments: 30 min

ANTICIPATED RESULTS

The described protocol allows overexpression of isotopically labeled proteins in human HEK293T cells and enables the acquisition of heteronuclear NMR experiments on living cells and cell lysates.

A number of proteins among those we tested reached expression levels suitable to be observed by in-cell NMR: SOD1 (both WT and fALS-related mutants), CCS, HAH1, Mia40, Cox17, profilin1 (with surface mutations), Grx1, Trx1 (**Fig. 3a**). The highest expression levels obtainable vary among proteins, ranging from ~150 μM (for SOD1) to ~10 μM (for Cox17), as measured from either Coomassie-stained SDS-PAGE or western blot (**Fig. 3b,c** and **Supplementary Fig. 2**). The expression levels can be lowered by decreasing the amount of DNA added in the transfection step, making possible to approach physiological concentrations, when the latter are compatible with NMR detectability (**Fig. 3d**). The expression levels obtained in each condition are very reproducible, falling within a 5% (s.d.) from the average calculated over several samples which were identically prepared in different years. Most of the proteins reported here can be clearly detected in the in-cell ^{1}H-^{15}N SOFAST-HMQC NMR spectra recorded on uniformly ^{15}N-labeled cell samples (**Fig. 4a-f**). These ^{1}H-^{15}N in-cell NMR spectra contain residue-level information of the protein conformational properties, and can be compared with the spectra acquired on the same proteins *in vitro* (**Fig. 4g-l**) to better characterize the intracellular species. Following this approach, different metallation states of SOD1 are easily identified: apo-SOD1, which is a partially disordered monomer (**Fig. 5a**, magenta), is readily converted to E,Zn-SOD1 (a well folded homodimer) when cells are treated with zinc (**Fig. 5a**, black).

Amino acid-selective labeling can be obtained by using home-made expression media supplemented with the desired labeled amino acid(s). For example, spectra can be obtained for samples grown with either [^{15}N]cysteine (**Fig. 5b**) or [^{13}C-methyl]methionine (**Fig. 5c**). The labeling efficiency when using [^{15}N]cysteine homemade media was estimated from crosspeak intensities as 75% ± 10% (s.d., calculated on 4 cysteine signals) of the efficiency obtained in the commercial [U-^{15}N] medium, while the intensity of the background signals was much lower (compare **Fig. 2a** and **5b**). Other amino acid-selective labeling schemes can be achieved in principle, however the amount of spurious isotopic incorporation is expected to vary with the nature of the amino acid. The [^{15}N]cysteine labeling is especially useful to investigate the redox state of the protein of interest, whenever the formation of disulfide bonds affects the amide chemical shifts of the involved cysteines. Indeed, the ^{1}H-^{15}N SOFAST-HMQC spectra acquired on HEK293T cells expressing [^{15}N]cysteine-labeled SOD1 showed that Cys146, which forms an intrasubunit disulfide bond with Cys57 in the mature enzyme, was reduced (**Fig. 5b**). The methionine labeling has the advantage of being relatively inexpensive, and allows the investigation of slow-tumbling macromolecules due to the favorably slower relaxation of the CH_3 moiety[39] (**Fig. 5c**).

ACKNOWLEDGEMENTS This work has been supported by the Access to Research Infrastructures activities in the Seventh Framework Programme of the European Commission (Bio-NMR contract 261863) and by Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI) and supported by national member subscriptions. Specifically, we thank the EU ESFRI Instruct Core Centre CERM-Italy.

AUTHOR CONTRIBUTIONS L. Barbieri, E.L. and L. Banci conceived the work and designed the experiments. L. Barbieri and E.L. performed the experiments. Specifically, L.

Barbieri seeded and transfected the cells, performed the SDS-PAGE and Western Blot analysis, and performed cell counts and viability tests; E.L. optimized the NMR experimental conditions, set up and acquired the NMR experiments, processed and analyzed the NMR data. L. Barbieri, E.L. and L. Banci wrote the manuscript.

COMPETING FINANCIAL INTERESTS The authors declare that they have no competing financial interests.

1. Reckel, S., Hänsel, R., Löhr, F. & Dötsch, V. In-cell NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **51,** 91–101 (2007).
2. Selenko, P. & Wagner, G. Looking into live cells with in-cell NMR spectroscopy. *J. Struct. Biol.* **158,** 244–253 (2007).
3. Inomata, K. *et al.* High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458,** 106–109 (2009).
4. Maldonado, A. Y., Burz, D. S. & Shekhtman, A. In-cell NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **59,** 197–212 (2011).
5. Banci, L. *et al.* Atomic-resolution monitoring of protein maturation in live human cells by NMR. *Nat. Chem. Biol.* **9,** 297–299 (2013).
6. Sakakibara, D. *et al.* Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458,** 102–105 (2009).
7. Waudby, C. A. *et al.* In-cell NMR characterization of the secondary structure populations of a disordered conformation of α-synuclein within E. coli cells. *PloS One* **8,** e72286 (2013).
8. Monteith, W. B. & Pielak, G. J. Residue level quantification of protein stability in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 11335–11340 (2014).
9. Burz, D. S., Dutta, K., Cowburn, D. & Shekhtman, A. Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat. Methods* **3,** 91–93 (2006).
10. Majumder, S. *et al.* Probing protein quinary interactions by in-cell nuclear magnetic resonance spectroscopy. *Biochemistry (Mosc.)* **54,** 2727–2738 (2015).
11. Monteith, W. B., Cohen, R. D., Smith, A. E., Guzman-Cisneros, E. & Pielak, G. J. Quinary structure modulates protein stability in cells. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 1739–1742 (2015).
12. Selenko, P., Serber, Z., Gadea, B., Ruderman, J. & Wagner, G. Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 11904–11909 (2006).
13. Ogino, S. *et al.* Observation of NMR signals from proteins introduced into living mammalian cells by reversible membrane permeabilization using a pore-forming toxin, streptolysin O. *J. Am. Chem. Soc.* **131,** 10834–10835 (2009).
14. Theillet, F.-X. *et al.* Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature* (2016). doi:10.1038/nature16531
15. Selenko, P. *et al.* In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat. Struct. Mol. Biol.* **15,** 321–329 (2008).
16. Bodart, J.-F. *et al.* NMR observation of Tau in Xenopus oocytes. *J. Magn. Reson.* **192,** 252–257 (2008).
17. Pielak, G. J. *et al.* Protein nuclear magnetic resonance under physiological conditions. *Biochemistry (Mosc.)* **48,** 226–234 (2009).
18. Hänsel, R., Foldynová-Trantírková, S., Dötsch, V. & Trantírek, L. Investigation of Quadruplex Structure Under Physiological Conditions Using In-Cell NMR. *Top. Curr. Chem.* (2012). doi:10.1007/128_2012_332
19. Danielsson, J. *et al.* Pruning the ALS-associated protein SOD1 for in-cell NMR. *J. Am. Chem. Soc.* **135,** 10266–10269 (2013).
20. Bertrand, K., Reverdatto, S., Burz, D. S., Zitomer, R. & Shekhtman, A. Structure of proteins in eukaryotic compartments. *J. Am. Chem. Soc.* **134,** 12798–12806 (2012).
21. Hamatsu, J. *et al.* High-resolution heteronuclear multidimensional NMR of proteins in living insect cells using a baculovirus protein expression system. *J. Am. Chem. Soc.* **135,** 1688–1691 (2013).
22. Banci, L., Barbieri, L., Luchinat, E. & Secci, E. Visualization of redox-controlled protein fold in living cells. *Chem. Biol.* **20,** 747–752 (2013).
23. Barbieri, L., Luchinat, E. & Banci, L. Structural insights of proteins in sub-cellular compartments: In-mitochondria NMR. *Biochim. Biophys. Acta* **1843,** 2492–2496 (2014).
24. Luchinat, E. *et al.* In-cell NMR reveals potential precursor of toxic species from SOD1 fALS mutants. *Nat. Commun.* **5,** 5502 (2014).
25. Barbieri, L., Luchinat, E. & Banci, L. Protein interaction patterns in different cellular environments are revealed by in-cell NMR. *Sci. Rep.* **5,** 14456 (2015).
26. Mercatelli, E., Barbieri, L., Luchinat, E. & Banci, L. Direct structural evidence of protein redox regulation obtained by in-cell NMR. *Biochim. Biophys. Acta* **1863,** 198–204 (2016).
27. Luchinat, E., Secci, E., Cencetti, F. & Bruni, P. Sequential protein expression and selective labeling for in-cell NMR in human cells. *Biochim. Biophys. Acta* **1860,** 527–533 (2016).
28. Banci, L. *et al.* Human superoxide dismutase 1 (hSOD1) maturation through interaction with human copper chaperone for SOD1 (hCCS). *Proc. Natl. Acad. Sci. U. S. A.* **109,** 13555–13560 (2012).
29. Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* **62,** 1243–1250 (2006).
30. Ye, J. *et al.* High-level protein expression in scalable CHO transient transfection. *Biotechnol. Bioeng.* **103,** 542–551 (2009).
31. Aruffo, A. Transient expression of proteins using COS cells. *Curr. Protoc. Mol. Biol. Ed. Frederick M Ausubel Al* **Chapter 16,** Unit 16.12 (2002).

32. Seiradake, E., Zhao, Y., Lu, W., Aricescu, A. R. & Jones, E. Y. Production of cell surface and secreted glycoproteins in mammalian cells. *Methods Mol. Biol. Clifton NJ* **1261,** 115–127 (2015).
33. Nettleship, J. E., Rahman-Huq, N. & Owens, R. J. The production of glycoproteins by transient expression in Mammalian cells. *Methods Mol. Biol. Clifton NJ* **498,** 245–263 (2009).
34. Kubo, S. *et al.* A gel-encapsulated bioreactor system for NMR studies of protein-protein interactions in living mammalian cells. *Angew. Chem. Int. Ed Engl.* (2012). doi:10.1002/anie.201207243
35. Wen, H., An, Y. J., Xu, W. J., Kang, K. W. & Park, S. Real-time monitoring of cancer cell metabolism and effects of an anticancer agent using 2D in-cell NMR spectroscopy. *Angew. Chem. Int. Ed Engl.* **54,** 5374–5377 (2015).
36. Hwang, T. L. & Shaka, A. J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson. A* **112,** 275–279 (1995).
37. Piotto, M., Saudek, V. & Sklenár, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2,** 661–665 (1992).
38. Schanda, P. & Brutscher, B. Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* **127,** 8014–8015 (2005).
39. Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E. & Kay, L. E. Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125,** 10420–10428 (2003).

Figures

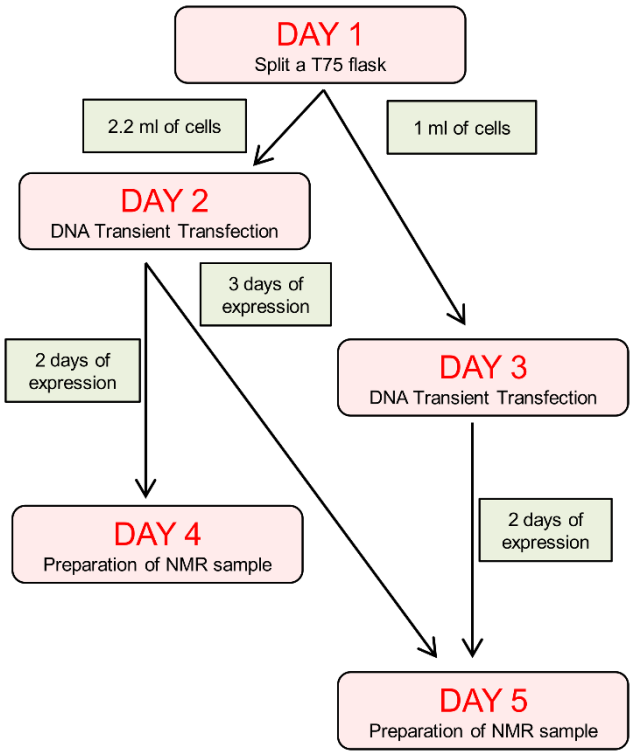


Figure 1 | Alternative timings for transfection and protein expression. Different timings of cell transfection and protein expression can be planned to fit the experimental needs and the NMR instrument schedule. Transfection can be performed either at day 2 or day 3 after seeding the cells at day 1. For transfection at day 2, seed 2.2 ml out of 10 ml of cell suspension (from a confluent T75 flask); for transfection at day 3, seed 1 ml out of 10 ml of cell suspension. The NMR sample will be prepared on different days, according to the planned protein expression time.

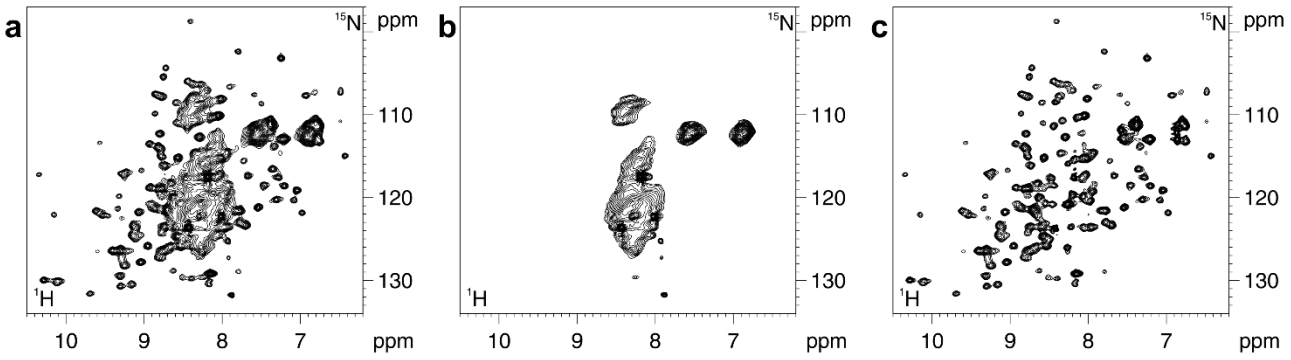


Figure 2 | Subtraction of 2D ^{1}H-^{15}N SOFAST-HMQC in-cell NMR spectra to eliminate the background signals. (a) NMR spectrum of cells expressing SOD1 in U-^{15}N-labeled medium. The amide signals of zinc-bound SOD1 are detected together with background signals (centred around 8.3 ppm) arising from other ^{15}N-labeled molecules. (b) NMR spectrum of a control cell sample transfected with an empty DNA vector. Only the background signals are detected. (c) NMR spectrum obtained by subtracting the background spectrum (b) to the original spectrum (a). Additional signals arising from SOD1 are now clearly identified in the central region.

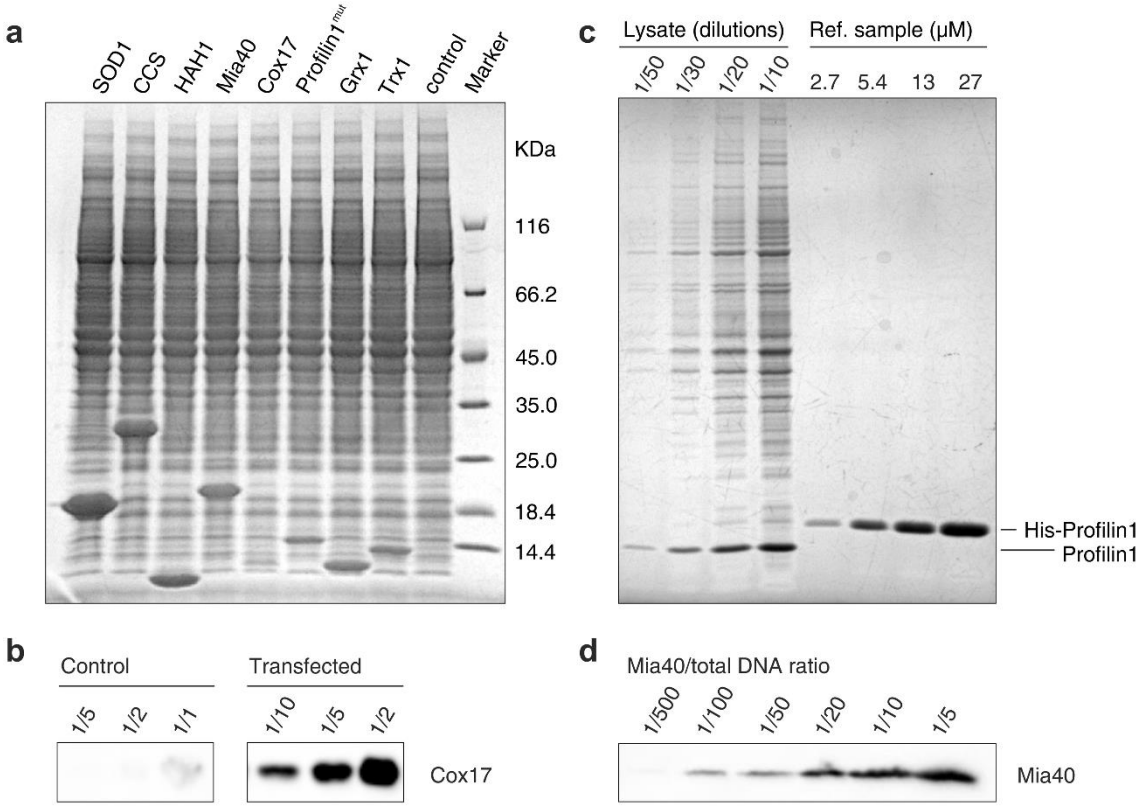


Figure 3 | Expression levels of different proteins transiently transfected in HEK293T cells. (a) SDS-PAGE analysis of cleared cell lysates obtained from cells overexpressing different proteins for 48 h. (b) Western Blot analysis of cleared cell lysates from cells transfected with empty vector (control) and Cox17 cDNA (transfected), loaded at different dilutions (vol/vol) and stained with anti-hCox17 antibody. (c) Quantification of overexpressed profilin1 by SDS-PAGE. Purified histidine-tagged profilin1 is used as a reference sample; both samples are loaded at different dilutions (vol/vol). (d) Dependence of the expression levels of Mia40 on the dilution of the corresponding DNA in the transfection mixture. Each cell lysate is loaded in the same amount in each lane. See also **Supplementary Fig. 2**.

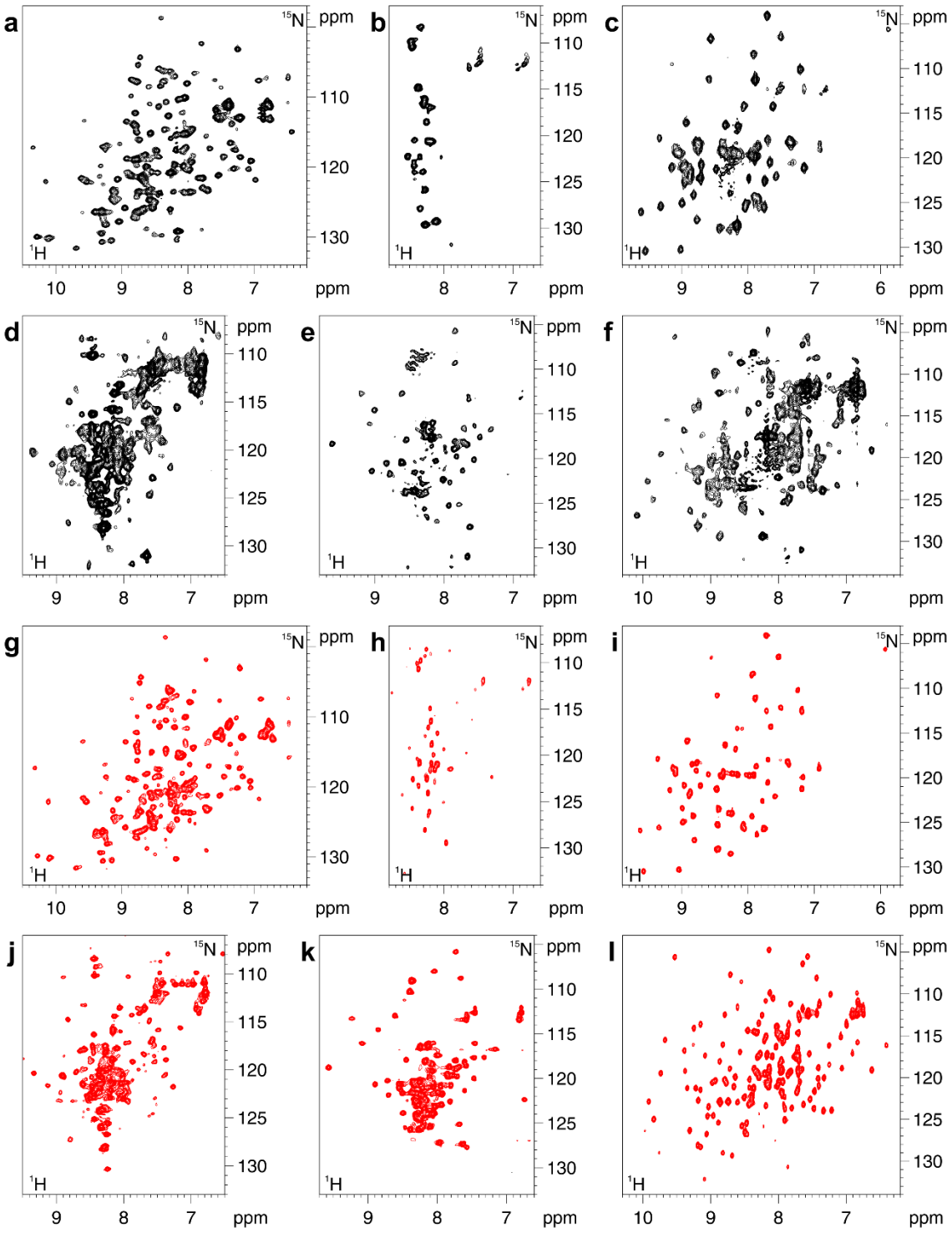


Figure 4 | ^{1}H-^{15}N in-cell NMR spectra (black) and the corresponding *in vitro* NMR spectra (red). (a-f) Background-subtracted ^{1}H-^{15}N SOFAST-HMQC in-cell NMR spectra recorded on cells expressing uniformly ^{15}N-labeled proteins. (a) SOD1; (b) CCS (only the C-terminal unfolded domain is detected); (c) HAH1; (d) Mia40; (e) Cox17; (f) profilin1 (with surface mutations). (g-l) ^{1}H-^{15}N NMR spectra recorded on purified ^{15}N-labeled samples of the same proteins as in (a-f). The spectrum of purified CCS (h) is a ^{1}H-^{15}N HSQC-TROSY, in which crosspeaks arising from the folded domains of CCS are also visible at lower intensity.

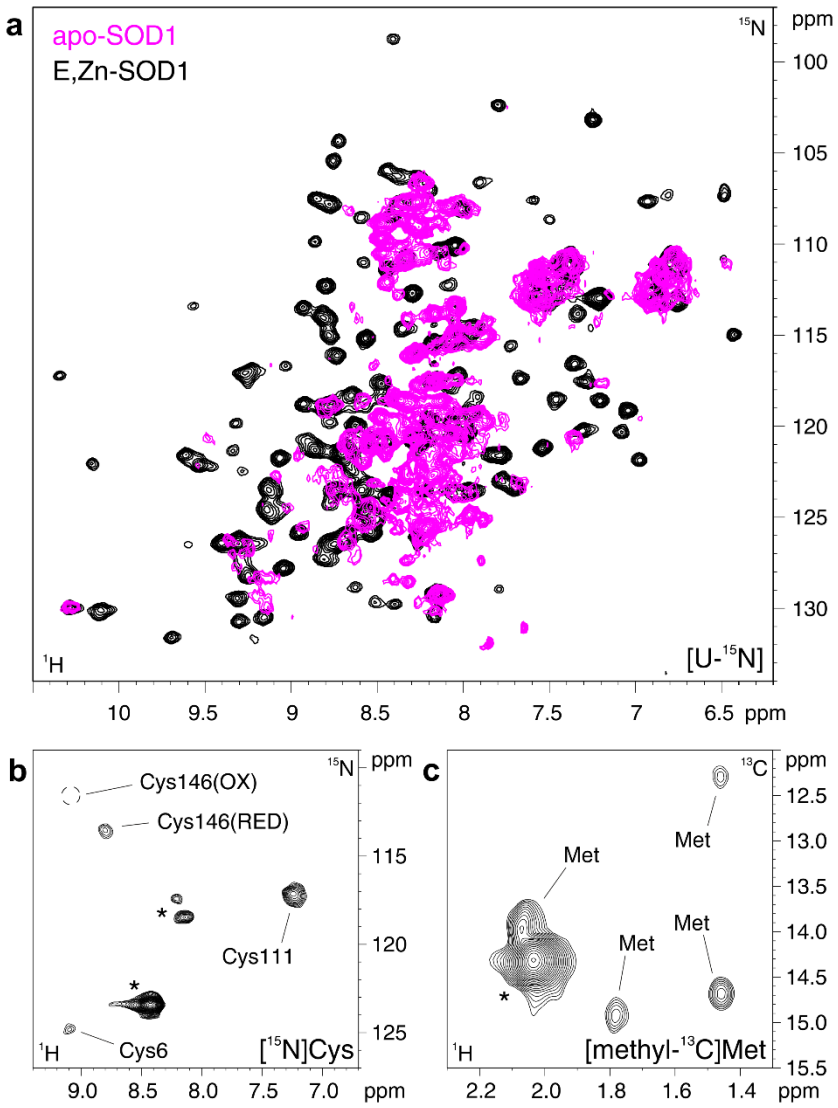
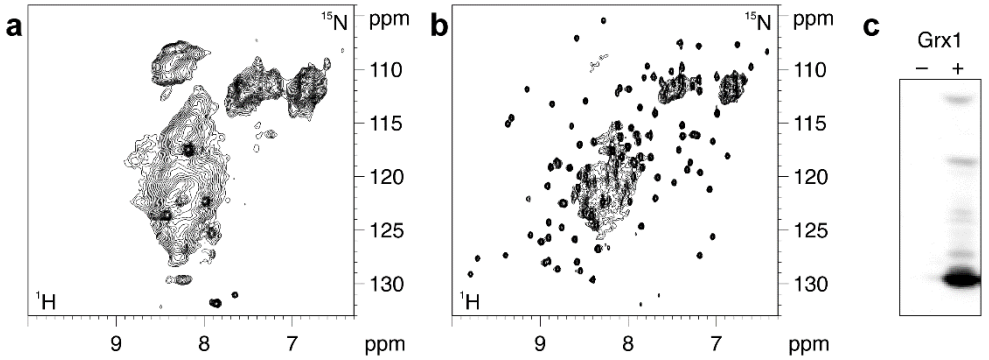
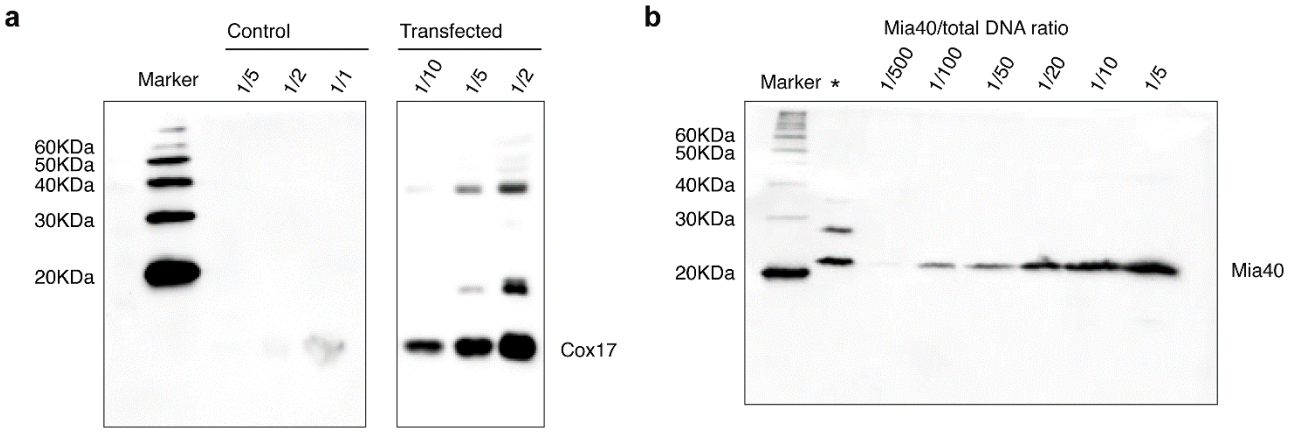


Figure 5 | (a) Overlay of ^{1}H-^{15}N SOFAST-HMQC in-cell NMR spectra recorded on cells expressing SOD1 without supplemented zinc (apo-SOD1, magenta) and in excess of zinc (E,Zn-SOD1, black). (b,c) 2D heteronuclear in-cell NMR spectra recorded on cells expressing selectively labeled proteins. (b) ^{1}H-^{15}N SOFAST-HMQC in-cell NMR spectrum of [^{15}N]cysteine-labeled SOD1. SOD1 cysteine crosspeaks are labeled. For the Cys146 crosspeak, the oxidation state is also indicated. (c) ^{1}H-^{13}C HSQC in-cell NMR spectrum of [methyl-^{13}C]methionine-labeled mutant profilin1. The methionine crosspeaks arising from the protein are indicated. Signals arising from cellular background are indicated with an asterisk.

Supplementary Figures



Supplementary Figure 1 | Grx1 is not detected in the in-cell NMR spectra. ^{1}H-^{15}N SOFAST-HMQC NMR spectra of (a) cells expressing [U-^{15}N]-labeled Grx1 and (b) the corresponding cell lysate. In the spectrum (a) the signals arising from Grx1 are not detected, due to the interaction with other cellular components. The two spectra were acquired with identical parameters; background subtraction was not performed. (c) Western Blot of the cell lysate analyzed in (b) (+) together with a control sample (–).



Supplementary Figure 2 | Western Blots. Full Western Blots corresponding to (a) **Fig. 3b** and (b) **Fig. 3d**. An unrelated sample is marked with an asterisk.