

In-Cell NMR in Human Cells: Direct Protein Expression Allows Structural Studies of Protein Folding and Maturation

Enrico Luchinat^{†,‡,§} and Lucia Banci^{*,†,§}

[†]Magnetic Resonance Center - CERM, University of Florence, 50019 Sesto Fiorentino, Italy

[‡]Department of Experimental and Clinical Biomedical Sciences “Mario Serio”, University of Florence, 50134 Florence, Italy

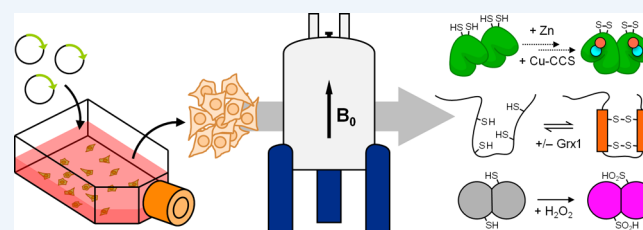
[§]Department of Chemistry, University of Florence, 50019 Sesto Fiorentino, Italy

CONSPECTUS: Cellular structural biology methods are needed to characterize biological processes at atomic resolution in the physiological environment of the cell. Toward this goal, solution in-cell NMR is a powerful approach because it provides structural and dynamic data on macromolecules inside living cells. Several approaches have been developed for in-cell NMR in cultured human cells, which are needed to study processes related to human diseases that rely on the delivery of exogenous macromolecules to the cells. Such strategies, however, may not be applicable to proteins that are sensitive to the external environment or prone to aggregate and can introduce artifacts during protein purification or delivery.

As a complementary approach, direct protein expression for in-cell NMR in human cells was developed. This strategy is especially useful when studying processes like protein folding, maturation, and post-translational modification, starting right after protein synthesis. Compared with the protein expression techniques in mammalian cells commonly used in cellular biology, the low sensitivity of NMR requires higher protein levels. Among the cell lines used for high-yield protein expression, the HEK293T cell line was chosen, as it can be efficiently transfected with a cost-effective reagent. A vector originally designed for secreted proteins allows high-level cytosolic protein expression. For isotopic labeling, commercially available or homemade labeled media are employed. Uniform or amino acid type-selective labeling strategies are possible. Protein expression can be targeted to specific organelles (e.g., mitochondria), allowing for in organello NMR applications. A variant of the approach was developed that allows the sequential expression of two or more proteins, with only one selectively labeled.

Protein expression in HEK293T cells was applied to recapitulate the maturation steps of intracellular superoxide dismutase 1 (SOD1) and to study the effect of mutations linked to familial amyotrophic lateral sclerosis (fALS) by in-cell NMR. Intracellular wild-type SOD1 spontaneously binds zinc, while it needs the copper chaperone for superoxide dismutase (CCS) for copper delivery and disulfide bond formation. Some fALS-linked mutations impair zinc binding and cause SOD1 to irreversibly unfold, likely forming the precursor of cytotoxic aggregates. The SOD-like domain of CCS acts as a molecular chaperone toward mutant SOD1, stabilizing its folding and allowing zinc binding and correct maturation. Changes in protein redox state distributions can also be investigated by in-cell NMR. Mitochondrial proteins require the redox-regulating partners glutaredoxin 1 (Grx1) and thioredoxin (Trx) to remain in the reduced, import-competent state in the cytosol, whereas SOD1 requires CCS for disulfide bond formation. In both cases, the proteins do not equilibrate with the cytosolic redox pool. Cysteine oxidation in response to oxidative stress can also be monitored.

In the near future, in-cell NMR in human cells will likely benefit from technological advancements in NMR hardware, the development of bioreactor systems for increased sample lifetime, the application of paramagnetic NMR to obtain structural restraints, and advanced tools for genome engineering and should be increasingly integrated with advanced cellular imaging techniques.



INTRODUCTION

Understanding biological processes requires a complete description of all of the involved molecules and their interactions at atomic resolution. The atomistic description of such processes is especially critical to develop novel chemicals and therapeutic protocols against human diseases. While this has always been the ultimate goal of structural biology, the classical approaches require each biomolecule to be isolated and analyzed far from its physiological context. Indeed, structural characterization is mainly performed by X-ray diffraction on

crystalline solids at cryogenic temperatures, whereas a minor fraction is performed by NMR spectroscopy on purified molecules in solution at room temperature and, recently, on vitrified samples by cryo-electron microscopy (cryo-EM). None of these environments comes even close to the actual physiological environment. In the last few decades, there have been increasing efforts to integrate data from classical

Received: March 30, 2018

Published: June 5, 2018

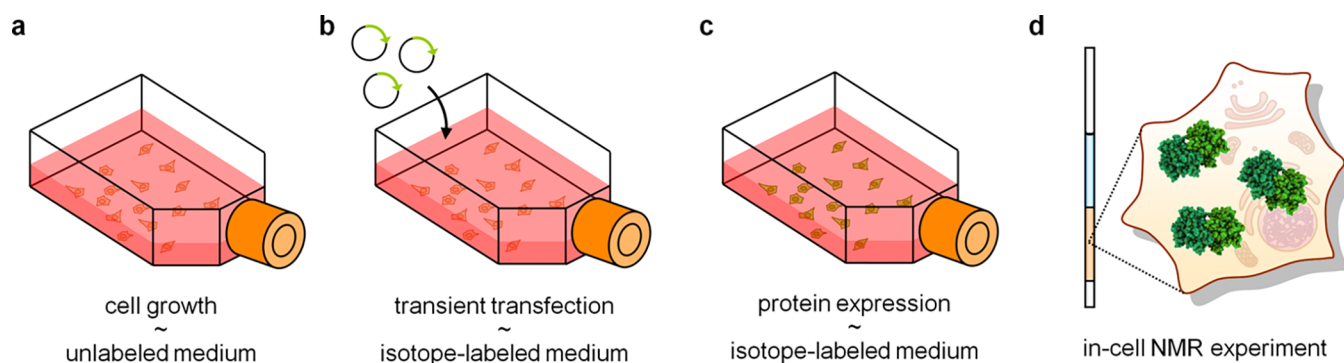


Figure 1. Protein expression in human cells for in-cell NMR. (a) HEK293T cells are seeded in a culture flask and grown in unlabeled medium. (b) The vector containing a gene of interest (green arrow) is delivered by transient transfection in labeled medium. (c) Protein expression is carried out in labeled medium. (d) Cells expressing the protein of interest (green) are collected and placed in a 3 mm Shigemi NMR tube for in-cell NMR analysis.

biochemistry, structural biology, and advanced cellular biology techniques (e.g., super-resolved microscopy) in order to build a more complete picture of the studied systems. However, such reductionist approaches are challenged by the need to make assumptions when working in vitro that may not be satisfied in vivo, while on the other hand the in vivo or in cellulo analysis is limited by the lack of atomic resolution. To overcome these limitations, cellular structural biology methods that would allow an atomistic description while preserving the physiological environment of the investigated system are being increasingly sought. In this context, NMR spectroscopy is the technique of choice, as it is nondestructive, works at physiological temperatures, and can monitor time-dependent phenomena on multiple time scales.

The first observation of an isotope-labeled protein by NMR in living bacterial cells marked the birth of an approach aptly named “in-cell NMR”.^{1,2} Clearly, observing macromolecules such as proteins in human cells is of critical importance when studying processes related to human diseases. Therefore, after some initial developments and applications in bacteria,³ the approach was extended to *Xenopus laevis* oocytes,⁴ the first eukaryotic model, and was eventually applied to observe a protein in human cells.⁵ That achievement provided a glimpse of the true potential of in-cell NMR and sparked the interest of the scientific community. In that work, proteins fused to the HIV-1 Tat cell-penetrating peptide were expressed in bacteria, purified, and subsequently delivered to the human cells by exploiting the mechanism of the viral peptide. Later, other approaches were developed for in-cell NMR and electron paramagnetic resonance (EPR) in human cells, all of which relied on alternative methods to deliver exogenous proteins, specifically treatment of the cells with the pore-forming toxin streptolysin O⁶ or permeabilization of the plasma membrane through hypotonic swelling⁷ or electroporation.⁸ These approaches have proven useful for observing intracellular protein^{5–9} and DNA¹⁰ signals in human cells. However, their application is often hampered by practical limitations, namely, the high concentration of external protein needed to obtain sufficiently high intracellular levels and the need to optimize the insertion technique to increase the internalization efficiency, which is highly dependent on the physicochemical properties of the investigated proteins. Furthermore, the insertion of human proteins purified from bacteria can introduce artifacts due to chemical modifications (or lack thereof) occurring in bacteria

or during purification, resulting in nonphysiological cofactor binding and/or redox states.

■ PROTEIN EXPRESSION IN HUMAN CELLS

The drawbacks of the protein insertion approaches are especially critical when monitoring some cellular processes, such as protein folding and maturation, cofactor binding, and changes in redox states, where the physiological relevance of the results requires each step to occur in the correct cellular environment, starting from protein synthesis. For these applications, the investigated proteins need to be expressed directly in the physiologically relevant host cells. While approaches for protein expression inside the cells analyzed by in-cell NMR were developed in yeast¹¹ and insect cells,¹² an approach relying on protein expression in human cells was missing, which would have been a step forward in the ability to study proteins related to human diseases in an environment as close as possible to the native cellular environment. In an effort to fill this gap and to further expand the range of applicability of in-cell NMR, our lab developed a method to express isotope-labeled proteins in human cells at levels suitable for NMR detection (Figure 1).^{13,14} Protein overexpression in mammalian cells is commonly used in cell biology, but the expression levels required for detection via Western blot or fluorescence microscopy are much lower than those needed for NMR. In the last few decades, several mammalian cell lines suitable for high-yield protein expression have been increasingly used both by the pharmaceutical industry for the production of biotherapeutic proteins and by scientists and biotech companies as a way to obtain challenging proteins. Seeking a cellular environment that would be as close to the native one as possible for studying human proteins, we focused on a human cell line, specifically the human embryonic kidney 293T (HEK293T) cell line.¹⁵ The method that we developed is based on an existing application of recombinant protein expression in HEK293T cells originally developed for the production of secreted glycosylated proteins for X-ray crystallography.¹⁶ In that work, transient transfection of adherent HEK293T cells was performed using branched polyethylenimine (PEI), a cationic polymer, as a cost-effective transfection reagent.¹⁷ PEI-mediated transfection ensures that high gene copy numbers are internalized in HEK293T cells with low toxicity. High-level protein expression is allowed by the pHLsec vector, which contains a strong constitutive synthetic promoter (CAG)¹⁸ followed by a secretion signal

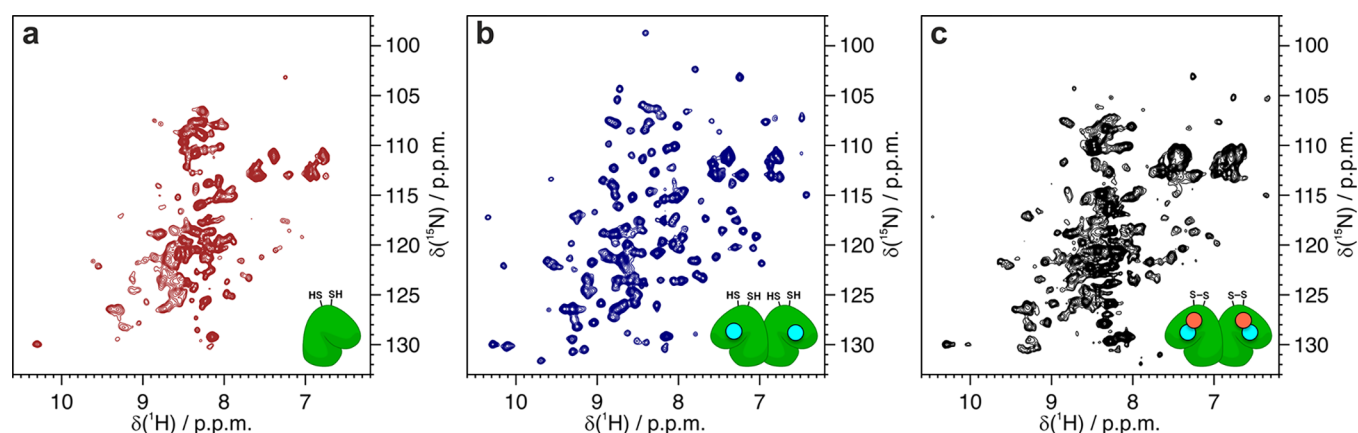


Figure 2. ^1H – ^{15}N NMR spectra of different SOD1 maturation states in human cells. (a) With a defect of zinc, SOD1 (green) is mostly in the apo, reduced state. (b) With zinc in excess, SOD1 dimerizes and binds one zinc ion (cyan) per monomer. (c) When CCS is coexpressed with zinc and copper in excess (salmon), the fully mature, disulfide-oxidized copper,zinc-SOD1 is formed. Also see ref 13.

sequence fused to the cDNA encoding the protein of interest.¹⁶ Because in-cell NMR aims to observe intracellular proteins, the secretion signal is removed from the original vector so that the protein of interest localizes in the cytoplasm. In the investigation of proteins by in-cell NMR, isotopic labeling is critical for two reasons: it allows heteronuclear multidimensional NMR experiments and acts as a filter to remove signals from the rest of the cell, so that only the labeled molecules are detected. In our protocol, isotopic labeling is performed by replacing the normal growth medium with an isotope-enriched one at the time of transfection. Commercially available media have been developed for uniform ^{15}N or ^{13}C , ^{15}N labeling in mammalian cells.¹⁹ Uniform labeling can also be achieved using custom-made media, e.g., those obtained from labeled algal autolysates.²⁰ Alternatively, amino acid type-selective labeling strategies are possible for certain amino acids, such as [^{15}N]cysteine and [*methyl*- ^{13}C]methionine.^{13,14} During protein expression phase in labeled media, other cellular components are partially labeled, resulting in the presence of cellular background signals in the NMR spectra. These background signals can be greatly reduced by subtracting NMR spectra acquired on a control sample of cells transfected with an empty vector, where protein expression did not occur.¹⁴ As is commonly the case with ectopic expression, this approach results in the cytosolic localization of the protein of interest and thus is best suited to investigate processes naturally occurring in this cellular compartment. Notably, cytosolic localization can also occur with proteins that are natively targeted to other compartments, such as mitochondrial proteins, likely because of the low efficiency of the native targeting sequence.²¹ However, we showed that fusion with a more efficient targeting sequence can effectively target the protein to the desired cellular compartment.²² In that work, a mitochondrial targeting sequence was used to target proteins in the mitochondrial intermembrane space (IMS). Intact mitochondria containing IMS-targeted labeled proteins were then isolated from HEK293T cells, and the protein signals could be detected in the resulting “in-mitochondria” NMR spectra. Therefore, in principle this approach can allow other in organello NMR studies to be performed by employing suitable targeting signals for different cellular compartments. The high gene copy number ensured by the employed transfection protocol allows two or more genes to be cotransfected, so that the encoded proteins are simultaneously expressed in the cells. However, for

proper characterization of protein–protein interactions by NMR, only one protein at a time should be labeled. For in-cell NMR purposes, this can be achieved by controlling the timing of expression so that the proteins are sequentially expressed and can be selectively labeled by appropriately switching the expression medium. This was achieved previously in bacteria by using expression vectors that could be independently induced.²³ For in-cell NMR in human cells, a sequential expression approach was developed in which a first gene is stably integrated into the host cell genome using an existing workflow.^{24,25} The obtained stable cell line is then cotransfected with the second gene and a mixture of vectors encoding small hairpin RNAs against the first gene. By appropriate timing of the incubation with labeled medium, only the second protein is selectively labeled while both proteins are present in the cells, allowing in principle the study of protein–protein interactions by NMR. To date, the direct expression approach for in-cell NMR has proven to be a versatile alternative to protein insertion, as shown already in other cellular systems.^{11,12} In human cells, several soluble proteins could be expressed at levels ranging from ~ 10 to ~ 150 μM ,¹⁴ which are sufficient for NMR detection provided that the proteins do not interact diffusely with other cellular components, causing excessive line broadening (a general limitation of solution in-cell NMR that is partly resolvable by introducing mutations on the protein surface²⁶). The approach has been applied to observe protein maturation and regulation processes, which are outlined below.

■ APPLICATIONS

Folding and Maturation of Superoxide Dismutase 1

Protein expression in human cells was first applied by our lab to investigate the folding and maturation pathway of wild-type (WT) human copper, zinc superoxide dismutase 1 (SOD1) by in-cell NMR.¹³ SOD1 is an evolutionarily conserved antioxidant enzyme that is present in most tissues at relatively high concentration, particularly in neuronal cells, and is localized in the cytoplasm, nucleus, and mitochondria. In order to reach the enzymatically active form, SOD1 needs to dimerize, bind one zinc ion and one copper ion per monomer, and form an intramolecular disulfide bond. In vivo, the latter two steps are catalyzed by the specific partner copper chaperone for superoxide dismutase (CCS).^{27,28} All of these events contribute

to increasing the stability of the mature protein. SOD1 has been linked to the onset of amyotrophic lateral sclerosis (ALS), a degenerative disease characterized by the death of motor neurons in the brain and spinal cord.²⁹ To date, more than 180 mutations scattered throughout the protein have been correlated to the onset of about 20% of familial ALS (fALS) cases.³⁰ Most mutations do not affect the enzymatic activity of the mature protein; instead, they have detrimental effects on various steps of the SOD1 maturation process through a toxic gain of function. As a consequence of impaired maturation, mutant SOD1 is less stable and more prone to misfold and eventually to form aggregates, which are a hallmark of the disease. The sequence of events leading to fully mature WT SOD1 was recapitulated by NMR in human cells by treating cells with different amounts of metal cofactors. With a defect of zinc, the majority of SOD1 is in the monomeric, partially unfolded apo state, with all of the cysteines in the reduced state (Figure 2a). With an excess of zinc, binding to the apo protein occurs spontaneously and quantitatively, and a zinc-bound dimeric reduced species is formed (Figure 2b). Unlike in vitro, a non-native form of SOD1 with a second zinc ion per monomer bound to the copper site is not observed in the cells, even with an excess of zinc. The higher selectivity exhibited in the cell is likely a consequence of the cellular zinc homeostasis. Intracellular copper is even more strictly regulated than zinc and has to be delivered to SOD1 by the metallochaperone CCS. Consistently, when copper is supplemented to the cells, it is bound only by a small fraction of SOD1, delivered either by means of the endogenous CCS, which cannot compensate for the higher levels of SOD1, or by a CCS-independent maturation pathway that has been reported for human SOD1.³¹ Cotransfection of both SOD1 and CCS cDNAs compensates for the increased levels of SOD1: upon copper treatment in the presence of higher levels of CCS, both copper binding and disulfide bond formation occur, and the mature form of SOD1 is observed (Figure 2c). Interestingly, CCS coexpression promotes the SOD1 disulfide bond formation also in the absence of copper, indicating that the metal transfer and redox reaction can be uncoupled in vivo under certain circumstances.

Given the importance of the SOD1 maturation pathway in the pathogenesis of fALS, we also investigated it for a set of fALS-linked SOD1 mutants.³² We focused on WT-like mutations, which do not perturb the metal binding sites of the protein or reduce the activity of the mature enzyme. Strikingly, a subset of the investigated SOD1 mutants failed to bind zinc in the cell even when zinc was available in excess and accumulated in the cytosol as unfolded species (Figure 3).³² Analysis of the cell lysates by NMR and size-exclusion chromatography revealed that these unfolded species are not oligomeric and are irreversibly formed, i.e., they cannot bind zinc even when zinc is supplemented after cell lysis. In vitro, the same mutant proteins are fully capable of binding zinc and exhibit a WT-like conformation both in the apo state and the zinc-bound state (Figure 3c), indicating that in the cell an irreversible unfolding/misfolding event had occurred prior to metal binding. These findings are consistent with a pathway of SOD1 misfolding and aggregation that starts from the apo reduced state, which is destabilized in vitro by fALS-linked mutations.³³ The unfolded species observed by NMR are therefore likely the precursors of aggregates, which have not yet formed because of the short experimental time (48 h of protein expression prior to NMR observation). Remarkably, coex-

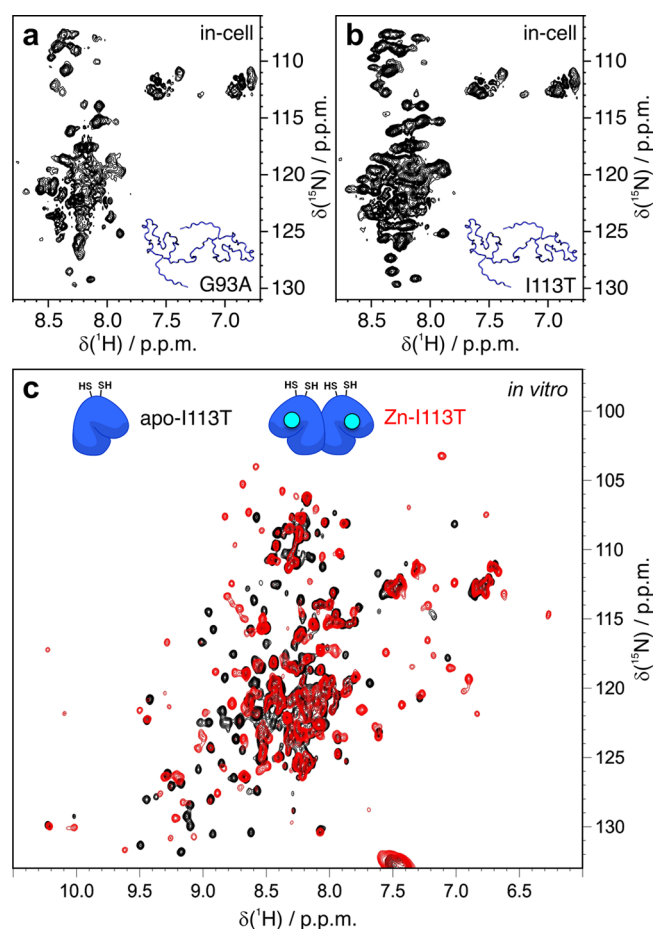


Figure 3. fALS-linked SOD1 mutations cause intracellular SOD1 unfolding and impaired zinc binding. (a, b) ^1H - ^{15}N in-cell NMR spectra of the SOD1 mutants (a) G93A and (b) I113T (blue), both present as unfolded apo species even with zinc in excess. (c) ^1H - ^{15}N in vitro NMR spectra of apo-I113T SOD1 (black spectrum) and Zn-I113T SOD1 (red spectrum) showing folded, WT-like conformations. Adapted with permission from ref 32. Copyright 2014 Springer Nature.

pression of CCS together with mutant SOD1 restored the correct maturation pathway, causing a drastic decrease in the amount of unfolded species and the appearance of the signals of the correctly folded SOD1.³² This effect of CCS, which is reportedly involved in the last steps of SOD1 maturation (copper binding and disulfide bond formation), prompted us to further investigate its role in the early steps (i.e., folding and zinc binding) of mutant SOD1 maturation. CCS consists of two globular domains: the N-terminal Atx1-like domain D1, which is responsible for copper(I) delivery, and the SOD-like domain D2, which forms a heterodimer with immature SOD1.³⁴ At the C-terminus, a disordered sequence (D3) catalyzes the formation of the SOD1 disulfide bond through a thiol-disulfide exchange mechanism, likely coupled to the copper(I) transfer.³⁵ We isolated the SOD1 recognition function of CCS by coexpressing only the second domain of CCS (D2) with the WT SOD1 and the fALS-linked mutants.³⁶ As expected, the CCS-dependent maturation does not occur without D1 and D3, and the SOD1-D2 heterodimer becomes a stable intermediate in the cell. Importantly, D2 in the cell acts as a molecular chaperone toward the destabilized SOD1 mutants. Indeed, D2 is able to interact with the apo form of SOD1, and

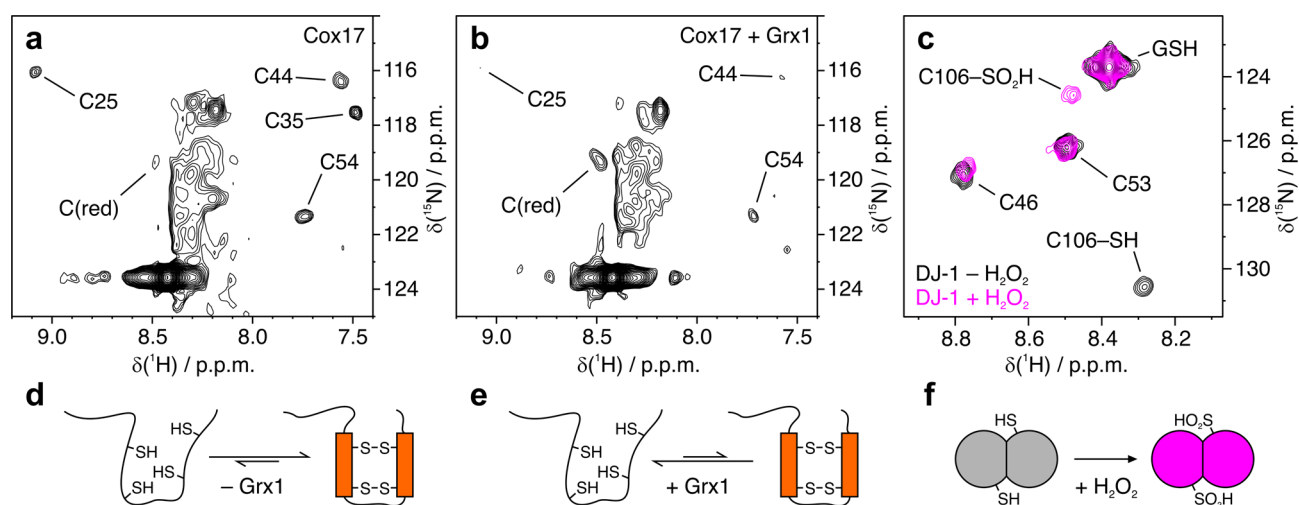


Figure 4. Protein redox state distribution monitored by in-cell NMR. (a–c) ^1H - ^{15}N in-cell NMR spectra of intracellular [^{15}N]cysteine-labeled Cox17 (a) without and (b) with coexpressed Grx1 and (c) [^{15}N]cysteine-labeled DJ-1 in control cells (black) and in cells treated with H_2O_2 (magenta). (d–f) Schematic illustrations of (d, e) thiol–disulfide regulation of Cox17 and (f) oxidation of DJ-1 C106 to sulfinic acid. Adapted with permission from (a, b) ref 39 and (c) ref 44. Copyright 2016 Elsevier and 2018 Springer Nature, respectively.

in doing so it stabilizes its fold and allows zinc binding, thereby preventing the irreversible misfolding of mutant SOD1. Such a novel molecular chaperone role of D2 highlights the importance of CCS in rescuing the correct folding of SOD1 and suggests future therapeutic strategies that potentiate this mechanism in fALS patients.

Protein Redox States and Regulation

As shown above, in-cell NMR can directly assess conformational changes caused by intracellular events such as metal binding and the formation of disulfide bonds. The function of many intracellular proteins is modulated by disulfide bond formation, which is regulated within different cellular compartments by specific redox partners. In-cell NMR can therefore provide unique insights into intracellular protein redox regulation at the molecular level thanks to the ability of NMR to identify directly the conformation of each redox state and to determine changes in the redox distribution as a function of intracellular partners or external stimuli. In this respect, direct protein expression is ideally suited for studying redox-sensitive proteins, which could be prone to artifacts if delivered from outside the cell. We first showed that the redox state of the mitochondrial protein Mia40 in the cytosol is regulated by the redox-regulating enzymes glutaredoxin 1 (Grx1) and thioredoxin (Trx).²¹ Mia40 is a redox chaperone of the IMS of mitochondria that is constituted by a coiled-coil helix, coiled-coil helix (CHCH) domain stabilized by two structural disulfide bonds.³⁷ In the IMS, Mia40 catalyzes the oxidative folding of other small mitochondrial proteins that harbor the same CHCH domain.³⁸ Like its substrates, Mia40 is synthesized in the cytosol and has to be imported into the mitochondria in the reduced, unfolded state. However, increased levels of Mia40 accumulate in the cytosol and are found mostly in the oxidized, folded state that is not import-competent. Coexpression of the thiol–disulfide-regulating proteins Grx1 and Trx shifts the redox distribution of cytosolic Mia40 toward the reduced state, indicating that proteins may not reach redox equilibrium with the environment unless the correct redox partners are present at sufficient levels. We further investigated the relationship between the redox properties of the environment and the redox state of cytosolic

proteins by in-cell NMR. The redox distribution of SOD1 and Cox17—a substrate of Mia40—was also analyzed in the cytosol of human cells (HEK293T), *Escherichia coli* (BL21), and a less reducing strain of *E. coli* (Origami B), which have different redox potentials as defined by the glutathione–glutathione disulfide couple.³⁹ Notably, the observed redox distribution clearly deviates from what would be expected at the redox equilibrium in each cellular environment. In human cells, coexpression of the known intracellular redox partners (Grx1/Trx for Mia40 and Cox17, CCS for SOD1) changes the redox distribution, either toward the equilibrium with glutathione (in the case of Mia40 and Cox17, Figure 4) or further away from it (in the case of SOD1). These findings support the notion that the redox state of most intracellular proteins is uncoupled from that of the glutathione pool and consequently that protein redox regulation needs to be kinetically controlled by specific partners, similar to how other post-translational modifications are regulated. Further evidence for kinetic control of the intracellular redox state was provided for the thioredoxin–thioredoxin reductase couple.⁴⁰

While under basal conditions the redox states of intracellular proteins are regulated by specific partners, they can dramatically change in response to external stimuli. In living organisms, an imbalance between the cellular production of reactive oxygen species and the antioxidant defense of the cell causes oxidative stress, which is involved in a plethora of physiological and pathological states, including aging, diabetes, and most degenerative diseases.⁴¹ At the molecular level, oxidative stress affects the redox states of many intracellular proteins, and in-cell NMR is the ideal methodology to observe such changes. We recently characterized the intracellular metal binding and redox state of DJ-1 under basal and stress conditions. DJ-1 (PARK7) is a ubiquitous protein involved in the cellular response against oxidative stress.⁴² DJ-1 has been associated with several pathologies, including cancer, Parkinson's disease, ALS, and ischemic injury. Despite the many roles attributed to DJ-1, ranging from proteasome regulation to chaperone and enzymatic activity, its precise function is not yet clear.⁴³ A redox-sensitive cysteine (C106) lies in the putative active site of the protein, for which different oxidation states have been

reported both in vitro and in vivo. C106 has been proposed to act as a ROS sensor, and its oxidation causes the loss of several functions of DJ-1. DJ-1 has also been reported to bind zinc and copper ions in its active site in vitro. The conformation of intracellular DJ-1 in differently treated cells was analyzed by comparing in-cell and in vitro NMR data in order to clarify the physiological state of the protein under basal and cell stress conditions.⁴⁴ In cells, no metal binding to DJ-1 occurred under any experimental conditions, consistent with the affinity of DJ-1 for zinc or copper being too low to compete against the metal buffering systems of the cell, whereas oxidative stress induced by H₂O₂ treatment caused the quantitative oxidation of the C106 thiol (–SH) to sulfinic acid (–SO₂H) (Figure 4c). No other previously reported oxidation states (i.e., sulfenic acid (–SOH) and sulfonic acid (–SO₃H)) were observed, as confirmed also by mass spectrometry, suggesting that in vivo DJ-1 may act as a redox sensor that switches between the C106 –SH and –SO₂H states.

FUTURE PERSPECTIVES

The applications described above show that protein expression in human cells can be successfully applied to investigate functional processes involving intracellular soluble proteins by solution NMR. Each specific case highlights the various types of structural perturbations that can be investigated: changes in the folding state (in the case of SOD1 and its mutants), thiol–disulfide redox regulation (SOD1, Mia40, and Cox17), metal binding (SOD1 and DJ-1), and response to oxidative stress (DJ-1). In all instances, the major strength of in-cell NMR is the ability to provide biologically meaningful structural information in the native cellular environment, which is then complemented by thorough in vitro NMR characterization of the various conformations. Overall, the in-cell NMR methodology has seen steady development in the last two decades, spanning several model organisms. In human cells, as an alternative to protein expression, insertion approaches have been successfully applied to study the effects of crowding on protein folding and conformation,⁸ post-translational modifications,⁴⁵ redox state,⁴⁰ and interactions with other cellular partners.⁹ Labeling tools have been developed, such as fluorine-containing amino acids to probe protein dynamics by in-cell ¹⁹F NMR⁴⁶ and lanthanide-based tags that can provide long-range spatial restraints by paramagnetic in-cell NMR.^{47,48} Furthermore, proteins in native membranes can be investigated by solid-state NMR⁴⁹ thanks to the sensitivity boost provided by dynamic nuclear polarization. In the near future, continuous technological developments will further increase the applicability of NMR to human cells. Progress in NMR hardware has constantly decreased the minimum protein concentration required. The development of bioreactors for high-field NMR spectrometers^{50,51} will increase the sample lifetime, allowing longer experimental times and, notably, continuous measurement of time-dependent processes. Finally, future approaches relying on protein expression will benefit from advanced tools for genome engineering, such as transposon-based systems,⁵² and finer control of the expression of two or more proteins with the use of inducible promoters. We believe that in-cell NMR will continue to be a powerful method to identify the true physiological states (versus the many possible nonphysiological states obtainable in vitro), acting as a guide to the in vitro analysis, and that it will be further integrated with other emerging cellular and structural techniques.^{53–55}

AUTHOR INFORMATION

Corresponding Author

*E-mail: banci@cerm.unifi.it.

ORCID

Enrico Luchinat: 0000-0003-4183-4311

Lucia Banci: 0000-0003-0562-5774

Notes

The authors declare no competing financial interest.

Biographies

Enrico Luchinat, born in Florence, Italy, on November 20, 1985, obtained a B.Sc. in Chemistry cum laude in 2007, an M.Sc. in Chemistry cum laude in 2009, and a Ph.D. in Mechanistic and Structural Systems Biology in 2013 at the University of Florence. He is currently a postdoctoral fellow at the University of Florence, where he is contributing to the development and application of methods for in-cell NMR in human cells.

Lucia Banci, born in Florence, Italy, on May 20, 1954, obtained a Doctorate in Chemistry cum laude in 1978. Since 1999 she has been Full Professor of Chemistry at the University of Florence. She is a cofounder and the former director of the Center of Magnetic Resonance (CERM) of the University of Florence and head of the Italian center of the ESFRI Landmark Research Infrastructure (Instruct-ERIC). She has been an EMBO member since 2012, and she received the IUPAC Distinguished Women in Chemistry Award in 2015 and the Instruct Bertini Award for Integrated Structural Biology in 2017. She has extensive expertise and has provided original contributions and breakthroughs in structural biology and biological NMR.

ACKNOWLEDGMENTS

The authors acknowledge Instruct ERIC, a Landmark ESFRI project, for supporting and stimulating the research described in the paper.

REFERENCES

- (1) Serber, Z.; Keatinge-Clay, A. T.; Ledwidge, R.; Kelly, A. E.; Miller, S. M.; Dötsch, V. High-Resolution Macromolecular NMR Spectroscopy inside Living Cells. *J. Am. Chem. Soc.* **2001**, *123*, 2446–2447.
- (2) Luchinat, E.; Banci, L. In-Cell NMR: A Topical Review. *IUCr* **2017**, *4*, 108–118.
- (3) Robinson, K. E.; Reardon, P. N.; Spicer, L. D. In-Cell NMR Spectroscopy in Escherichia Coli. *Methods Mol. Biol.* **2012**, *831*, 261–277.
- (4) 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.* **2006**, *103*, 11904–11909.
- (5) Inomata, K.; Ohno, A.; Tochio, H.; Isogai, S.; Tenno, T.; Nakase, I.; Takeuchi, T.; Futaki, S.; Ito, Y.; Hiroaki, H.; Shirakawa, M. High-Resolution Multi-Dimensional NMR Spectroscopy of Proteins in Human Cells. *Nature* **2009**, *458*, 106–109.
- (6) Ogino, S.; Kubo, S.; Umamoto, R.; Huang, S.; Nishida, N.; Shimada, I. 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.* **2009**, *131*, 10834–10835.
- (7) Martorana, A.; Bellapadrone, G.; Feintuch, A.; Di Gregorio, E.; Aime, S.; Goldfarb, D. Probing Protein Conformation in Cells by EPR Distance Measurements Using Gd³⁺ Spin Labeling. *J. Am. Chem. Soc.* **2014**, *136*, 13458–13465.

- (8) Theillet, F.-X.; Binolfi, A.; Bekei, B.; Martorana, A.; Rose, H. M.; Stuver, M.; Verzini, S.; Lorenz, D.; van Rossum, M.; Goldfarb, D.; Selenko, P. Structural Disorder of Monomeric α -Synuclein Persists in Mammalian Cells. *Nature* **2016**, *530*, 45–50.
- (9) Majumder, S.; Xue, J.; DeMott, C. M.; Reverdatto, S.; Burz, D. S.; Shekhtman, A. Probing Protein Quinary Interactions by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2015**, *54*, 2727–2738.
- (10) Dzatko, S.; Krafcikova, M.; Hänsel-Hertsch, R.; Fessl, T.; Fiala, R.; Loja, T.; Krafcik, D.; Mergny, J.-L.; Foldynova-Trantirkova, S.; Trantirek, L. Evaluation of the Stability of DNA I-Motifs in the Nuclei of Living Mammalian Cells. *Angew. Chem., Int. Ed.* **2018**, *57*, 2165–2169.
- (11) Bertrand, K.; Reverdatto, S.; Burz, D. S.; Zitomer, R.; Shekhtman, A. Structure of Proteins in Eukaryotic Compartments. *J. Am. Chem. Soc.* **2012**, *134*, 12798–12806.
- (12) Hamatsu, J.; O'Donovan, D.; Tanaka, T.; Shirai, T.; Hourai, Y.; Mikawa, T.; Ikeya, T.; Mishima, M.; Boucher, W.; Smith, B. O.; Laue, E. D.; Shirakawa, M.; Ito, Y. High-Resolution Heteronuclear Multidimensional NMR of Proteins in Living Insect Cells Using a Baculovirus Protein Expression System. *J. Am. Chem. Soc.* **2013**, *135*, 1688–1691.
- (13) Banci, L.; Barbieri, L.; Bertini, I.; Luchinat, E.; Secci, E.; Zhao, Y.; Aricescu, A. R. Atomic-Resolution Monitoring of Protein Maturation in Live Human Cells by NMR. *Nat. Chem. Biol.* **2013**, *9*, 297–299.
- (14) Barbieri, L.; Luchinat, E.; Banci, L. Characterization of Proteins by In-Cell NMR Spectroscopy in Cultured Mammalian Cells. *Nat. Protoc.* **2016**, *11*, 1101–1111.
- (15) Dumont, J.; Ewart, D.; Mei, B.; Estes, S.; Kshirsagar, R. Human Cell Lines for Biopharmaceutical Manufacturing: History, Status, and Future Perspectives. *Crit. Rev. Biotechnol.* **2016**, *36*, 1110–1122.
- (16) Aricescu, A. R.; Lu, W.; Jones, E. Y. A Time- and Cost-Efficient System for High-Level Protein Production in Mammalian Cells. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2006**, *62*, 1243–1250.
- (17) Durocher, Y.; Perret, S.; Kamen, A. High-Level and High-Throughput Recombinant Protein Production by Transient Transfection of Suspension-Growing Human 293-EBNA1 Cells. *Nucleic Acids Res.* **2002**, *30*, e9.
- (18) Niwa, H.; Yamamura, K.-i.; Miyazaki, J.-i. Efficient Selection for High-Expression Transfectants with a Novel Eukaryotic Vector. *Gene* **1991**, *108*, 193–199.
- (19) Dutta, A.; Saxena, K.; Schwalbe, H.; Klein-Seetharaman, J. Isotope Labeling in Mammalian Cells. *Methods Mol. Biol.* **2012**, *831*, 55–69.
- (20) Fuccio, C.; Luchinat, E.; Barbieri, L.; Neri, S.; Fragai, M. Algal Autolysate Medium to Label Proteins for NMR in Mammalian Cells. *J. Biomol. NMR* **2016**, *64*, 275–280.
- (21) Banci, L.; Barbieri, L.; Luchinat, E.; Secci, E. Visualization of Redox-Controlled Protein Fold in Living Cells. *Chem. Biol.* **2013**, *20*, 747–752.
- (22) Barbieri, L.; Luchinat, E.; Banci, L. Structural Insights of Proteins in Sub-Cellular Compartments: In-Mitochondria NMR. *Biochim. Biophys. Acta, Mol. Cell Res.* **2014**, *1843*, 2492–2496.
- (23) Burz, D. S.; Dutta, K.; Cowburn, D.; Shekhtman, A. Mapping Structural Interactions Using In-Cell NMR Spectroscopy (STINT-NMR). *Nat. Methods* **2006**, *3*, 91–93.
- (24) 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.* **2015**, *1261*, 115–127.
- (25) Luchinat, E.; Secci, E.; Cencetti, F.; Bruni, P. Sequential Protein Expression and Selective Labeling for In-Cell NMR in Human Cells. *Biochim. Biophys. Acta, Gen. Subj.* **2016**, *1860*, 527–533.
- (26) Barbieri, L.; Luchinat, E.; Banci, L. Protein Interaction Patterns in Different Cellular Environments Are Revealed by In-Cell NMR. *Sci. Rep.* **2015**, *5*, 14456.
- (27) Culotta, V. C.; Klomp, L. W.; Strain, J.; Casareno, R. L.; Krems, B.; Gitlin, J. D. The Copper Chaperone for Superoxide Dismutase. *J. Biol. Chem.* **1997**, *272*, 23469–23472.
- (28) Banci, L.; Bertini, I.; Cantini, F.; Kozyreva, T.; Massagni, C.; Palumaa, P.; Rubino, J. T.; Zovo, K. Human Superoxide Dismutase 1 (HSOD1) Maturation through Interaction with Human Copper Chaperone for SOD1 (HCCS). *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 13555–13560.
- (29) Rosen, D. R.; Siddique, T.; Patterson, D.; Figlewicz, D. A.; Sapp, P.; Hentati, A.; Donaldson, D.; Goto, J.; O'Regan, J. P.; Deng, H. X.; et al. Mutations in Cu/Zn Superoxide Dismutase Gene Are Associated with Familial Amyotrophic Lateral Sclerosis. *Nature* **1993**, *362*, 59–62.
- (30) Abel, O.; Powell, J. F.; Andersen, P. M.; Al-Chalabi, A. ALSod: A User-Friendly Online Bioinformatics Tool for Amyotrophic Lateral Sclerosis Genetics. *Hum. Mutat.* **2012**, *33*, 1345–1351.
- (31) Carroll, M. C.; Girouard, J. B.; Ulloa, J. L.; Subramaniam, J. R.; Wong, P. C.; Valentine, J. S.; Culotta, V. C. Mechanisms for Activating Cu- and Zn-Containing Superoxide Dismutase in the Absence of the CCS Cu Chaperone. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5964–5969.
- (32) Luchinat, E.; Barbieri, L.; Rubino, J. T.; Kozyreva, T.; Cantini, F.; Banci, L. In-Cell NMR Reveals Potential Precursor of Toxic Species from SOD1 FALS Mutants. *Nat. Commun.* **2014**, *5*, 5502.
- (33) Lindberg, M. J.; Tibell, L.; Oliveberg, M. Common Denominator of Cu/Zn Superoxide Dismutase Mutants Associated with Amyotrophic Lateral Sclerosis: Decreased Stability of the Apo State. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 16607–16612.
- (34) Lamb, A. L.; Torres, A. S.; O'Halloran, T. V.; Rosenzweig, A. C. Heterodimeric Structure of Superoxide Dismutase in Complex with Its Metallochaperone. *Nat. Struct. Biol.* **2001**, *8*, 751–755.
- (35) Fetherolf, M. M.; Boyd, S. D.; Taylor, A. B.; Kim, H. J.; Wohlschlegel, J. A.; Blackburn, N. J.; Hart, P. J.; Winge, D. R.; Winkler, D. D. Copper-Zinc Superoxide Dismutase Is Activated through a Sulfenic Acid Intermediate at a Copper Ion Entry Site. *J. Biol. Chem.* **2017**, *292*, 12025–12040.
- (36) Luchinat, E.; Barbieri, L.; Banci, L. A Molecular Chaperone Activity of CCS Restores the Maturation of SOD1 FALS Mutants. *Sci. Rep.* **2017**, *7*, 17433.
- (37) Banci, L.; Bertini, I.; Cefaro, C.; Ciofi-Baffoni, S.; Gallo, A.; Martinelli, M.; Sideris, D. P.; Katrakili, N.; Tokatlidis, K. MIA40 Is an Oxidoreductase That Catalyzes Oxidative Protein Folding in Mitochondria. *Nat. Struct. Mol. Biol.* **2009**, *16*, 198–206.
- (38) Banci, L.; Bertini, I.; Cefaro, C.; Cenacchi, L.; Ciofi-Baffoni, S.; Felli, I. C.; Gallo, A.; Gonnelli, L.; Luchinat, E.; Sideris, D.; Tokatlidis, K. Molecular Chaperone Function of Mia40 Triggers Consecutive Induced Folding Steps of the Substrate in Mitochondrial Protein Import. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 20190–20195.
- (39) Mercatelli, E.; Barbieri, L.; Luchinat, E.; Banci, L. Direct Structural Evidence of Protein Redox Regulation Obtained by In-Cell NMR. *Biochim. Biophys. Acta, Mol. Cell Res.* **2016**, *1863*, 198–204.
- (40) Mochizuki, A.; Saso, A.; Zhao, Q.; Kubo, S.; Nishida, N.; Shimada, I. Balanced Regulation of Redox Status of Intracellular Thioredoxin Revealed by In-Cell NMR. *J. Am. Chem. Soc.* **2018**, *140*, 3784–3790.
- (41) Uttara, B.; Singh, A. V.; Zamboni, P.; Mahajan, R. Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. *Curr. Neuropharmacol.* **2009**, *7*, 65–74.
- (42) Biosia, A.; Sandrelli, F.; Beltrami, M.; Greggio, E.; Bubacco, L.; Bisaglia, M. Recent Findings on the Physiological Function of DJ-1: Beyond Parkinson's Disease. *Neurobiol. Dis.* **2017**, *108*, 65–72.
- (43) Richarme, G.; Liu, C.; Mihoub, M.; Abdallah, J.; Leger, T.; Joly, N.; Liebart, J.-C.; Jurkunas, U. V.; Nadal, M.; Boulou, P.; Dairou, J.; Lamouri, A. Guanine Glycation Repair by DJ-1/Park7 and Its Bacterial Homologs. *Science* **2017**, *357*, 208–211.
- (44) Barbieri, L.; Luchinat, E.; Banci, L. Intracellular Metal Binding and Redox Behavior of Human DJ-1. *J. Biol. Inorg. Chem.* **2018**, *23*, 61–69.

(45) Binolfi, A.; Limatola, A.; Verzini, S.; Kosten, J.; Theillet, F.-X.; Rose, H. M.; Bekei, B.; Stuiiver, M.; van Rossum, M.; Selenko, P. Intracellular Repair of Oxidation-Damaged α -Synuclein Fails to Target C-Terminal Modification Sites. *Nat. Commun.* **2016**, *7*, 10251.

(46) Li, C.; Wang, G.-F.; Wang, Y.; Creager-Allen, R.; Lutz, E. A.; Scronce, H.; Slade, K. M.; Ruf, R. A. S.; Mehl, R. A.; Pielak, G. J. Protein (19)F NMR in Escherichia Coli. *J. Am. Chem. Soc.* **2010**, *132*, 321–327.

(47) Müntener, T.; Häussinger, D.; Selenko, P.; Theillet, F.-X. In-Cell Protein Structures from 2D NMR Experiments. *J. Phys. Chem. Lett.* **2016**, *7*, 2821–2825.

(48) Pan, B.-B.; Yang, F.; Ye, Y.; Wu, Q.; Li, C.; Huber, T.; Su, X.-C. 3D Structure Determination of a Protein in Living Cells Using Paramagnetic NMR Spectroscopy. *Chem. Commun.* **2016**, *52*, 10237–10240.

(49) Kaplan, M.; Narasimhan, S.; de Heus, C.; Mance, D.; van Doorn, S.; Houben, K.; Popov-Celeketić, D.; Damman, R.; Katrukha, E. A.; Jain, P.; Geerts, W. J. C.; Heck, A. J. R.; Folkers, G. E.; Kapitein, L. C.; Lemeer, S.; van Bergen en Henegouwen, P. M. P.; Baldus, M. EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. *Cell* **2016**, *167*, 1241–1251.

(50) Kubo, S.; Nishida, N.; Udagawa, Y.; Takarada, O.; Ogino, S.; Shimada, I. A Gel-Encapsulated Bioreactor System for NMR Studies of Protein-Protein Interactions in Living Mammalian Cells. *Angew. Chem., Int. Ed.* **2013**, *52*, 1208–1211.

(51) Breindel, L.; DeMott, C.; Burz, D. S.; Shekhtman, A. Real-Time In-Cell Nuclear Magnetic Resonance: Ribosome-Targeted Antibiotics Modulate Quinary Protein Interactions. *Biochemistry* **2018**, *57*, 540–546.

(52) Li, X.; Burnight, E. R.; Cooney, A. L.; Malani, N.; Brady, T.; Sander, J. D.; Staber, J.; Wheelan, S. J.; Joung, J. K.; McCray, P. B.; Bushman, F. D.; Sinn, P. L.; Craig, N. L. PiggyBac Transposase Tools for Genome Engineering. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E2279–E2287.

(53) Luchinat, E.; Gianoncelli, A.; Mello, T.; Galli, A.; Banci, L. Combining In-Cell NMR and X-Ray Fluorescence Microscopy to Reveal the Intracellular Maturation States of Human Superoxide Dismutase 1. *Chem. Commun.* **2015**, *51*, 584–587.

(54) Mitri, E.; Barbieri, L.; Vaccari, L.; Luchinat, E. ^{15}N Isotopic Labelling for In-Cell Protein Studies by NMR Spectroscopy and Single-Cell IR Synchrotron Radiation FTIR Microscopy: A Correlative Study. *Analyst* **2018**, *143*, 1171–1181.

(55) Baker, L. A.; Sinnige, T.; Schellenberger, P.; de Keyser, J.; Siebert, C. A.; Driessen, A. J. M.; Baldus, M.; Grünewald, K. Combined ^1H -Detected Solid-State NMR Spectroscopy and Electron Cryotomography To Study Membrane Proteins across Resolutions in Native Environments. *Structure* **2018**, *26*, 161–170.