



## King's Research Portal

DOI:

[10.1016/j.abb.2017.02.008](https://doi.org/10.1016/j.abb.2017.02.008)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Pastore, A., & Temussi, P. A. (2017). The Emperor's new clothes: Myths and truths of in-cell NMR. *Archives of biochemistry and biophysics*. <https://doi.org/10.1016/j.abb.2017.02.008>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# Accepted Manuscript

The Emperor's new clothes: Myths and truths of in-cell NMR

Annalisa Pastore, Piero Andrea Temussi

PII: S0003-9861(17)30006-1

DOI: [10.1016/j.abb.2017.02.008](https://doi.org/10.1016/j.abb.2017.02.008)

Reference: YABBI 7441

To appear in: *Archives of Biochemistry and Biophysics*

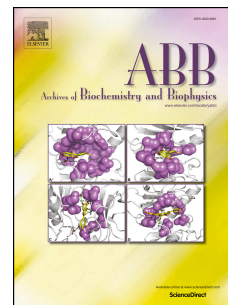
Received Date: 2 January 2017

Revised Date: 24 February 2017

Accepted Date: 27 February 2017

Please cite this article as: A. Pastore, P.A. Temussi, The Emperor's new clothes: Myths and truths of in-cell NMR, *Archives of Biochemistry and Biophysics* (2017), doi: 10.1016/j.abb.2017.02.008.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



## The Emperor's New Clothes: myths and truths of in-cell NMR

Annalisa Pastore<sup>1,2\*</sup>, Piero Andrea Temussi<sup>1,3</sup>

<sup>1</sup>The Wohl Institute, King's College London, 5 Cutcombe Rd, London SE5 9RT (UK)

<sup>2</sup>University of Pavia, Department of Molecular Medicine, Pavia (Italy)

<sup>3</sup>University of Naples "Federico II", Department of Chemical Sciences, Naples (Italy)

\*Corresponding author

**Running title:** In-cell NMR

**Keywords:** cellular environment, confinement, molecular crowding, NMR

### Abstract

In-cell NMR is a technique developed to study the structure and dynamical behavior of biological macromolecules in their natural environment, circumventing all isolation and purification steps. In principle, the potentialities of the technique are enormous, not only for the possibility of bypassing all purification steps but, even more importantly, for the wealth of information that can be gained from directly monitoring interactions among biological macromolecules in a natural cell. Here, we review critically the promises, successes and limits of this technique as it stands now. Interestingly, many of the problems of NMR in bacterial cells stem from the artificially high concentration of the protein under study whose overexpression is anyway necessary to select it from the background. This has, as a consequence, that when overexpressed, most globular proteins, do not show an NMR spectrum, limiting the applicability of the technique to intrinsically unfolded or specifically behaving proteins. The outlook for in-cell NMR of eukaryotic cells is more promising and is possibly the most attracting aspect for the future.

## Introduction

The environment generally used in *in vitro* studies of bio-macromolecules is a dilute solution of the highly-purified macromolecule of interest, which contains only water, buffer and salts. Thanks to the pioneering work of Minton [1], it is now widely accepted that the dilute solution paradigm does not reproduce cell environments faithfully. The cell interior contains from 20 to 40% (in weight) of very different macromolecules [2, 3]. Accordingly, the cell medium is described as being both crowded and confining. Although the two terms are often used as synonyms, they are not equivalent: crowding refers to a dynamic situation whereas confinement refers to a static one, assured by compartments [1, 4, 5].

Crowding may play an important role if one considers that the distribution of proteins in cells is far from uniform. Many proteins are part of big complexes or occupy organelles whose volume is a tiny portion of the whole cell [6]. Confinement (in the cell volume or in smaller organelles) is another means to exclude the volume available to macromolecular solutes. Volume exclusion related to crowding and confinement has important consequences for protein stability and reactivity. Combinations of dynamic (crowding) and static (confinement) volume exclusion can be characteristic of many processes in the living environment. In addition to the constraints imposed by crowding and confinement, proteins in living cells experience both specific and unspecific interactions with other macromolecules. Studies mimicking cell conditions are plagued by difficulties, mainly because the choice of appropriate crowders and of their concentrations is far from obvious [7].

The natural answer to the inadequacies of biophysical studies in dilute *in vitro* environments would be to perform these studies observing the macromolecule of interest directly inside cells. Such an approach is possible in principle and has in fact been pursued by several researchers, but it still presents difficulties. The best available technique is probably NMR spectroscopy, usually combined with selective labeling of one of the atomic species present in proteins under study, notably  $^{15}\text{N}$  labeling, and overexpression of the same proteins in a suitable cell, frequently an *E. coli*. All other molecules become effectively invisible if one uses pulse sequences tuned on the resonance of  $^{15}\text{N}$ .

In-cell NMR is distinct from the wealth of *in vivo* NMR studies that had preceded it, particularly metabolomics [8] or imaging studies. According to the researchers that introduced it, the goal of this technique is to study the structure and dynamical behavior of biological macromolecules in their natural milieu [9]. The potentialities of the technique were impressive: it seemed possible to determine protein structure in solution directly in cell, bypassing all purification steps and, even

more importantly, it promised to allow the study of interactions among biological macromolecules in a natural cell environment.

However, fifteen years after its first introduction, it is apparent that in-cell NMR seems to be possible only in specific cases: it is now clear that the behavior of the protein at the center of the first pioneering study of Serber et al. [10], namely NmerA, is an exception rather than the rule. Following the original study on NmerA, a few other proteins could be observed directly in prokaryotic cells [11]; notable among them is the B domain of G protein (GB1) and the putative heavy-metal binding protein TTHA1718, the only protein whose structure was solved by in-cell NMR [12]. A few in-cell NMR studies did find small differences between in-vitro and in-cell NMR parameters and/or evidence of structuring of intrinsically unfolded proteins in cell. However, when overexpressed in bacteria, most folded proteins, do not show an NMR spectrum [13]. The absence of the NMR spectrum has been attributed to several causes, such as the marked crowding of the bacterial cytoplasm and protein stickiness, but many of the problems of NMR in bacterial cells stem from the artificially high concentration of the protein under study, whose overexpression is anyway necessary to select it from the background. The outlook for in-cell NMR of eukaryotic cells is more promising.

There are several excellent reviews on in-cell NMR [11, 13, 14-24]. Therefore here, after revisiting the historical aspects of in-cell NMR, we mainly review its most recent developments, concentrating on the unsolved problems and the perspectives of the method. We are aware that, in doing so, we may have inevitably omitted references to many valuable papers and apologize for these unintentional omissions.

## **How to do in-cell NMR: Technical aspects**

### *Expressing proteins in Prokaryotes*

The branch of NMR spectroscopy generally called in-cell NMR was based on overexpression of the labeled protein in *E. coli* cells [10, 14]. To distinguish the NMR spectrum of the bio-macromolecules under study from generic signals of cellular components it is essential to selectively label the macromolecules. The isotopes generally used are  $^{15}\text{N}$  or  $^{13}\text{C}$ , but  $^{19}\text{F}$  is also increasingly being used [25, 26]. In principle, overexpressing labeled macromolecules inside cells requires an isotopically labeled medium. Although this procedure implies the risk of strong background signals, detailed investigations from the group that originated in-cell NMR showed that  $^{15}\text{N}$  labeling generally leads to just a few background peaks, even when the cells are grown from the beginning in labeled media [14]. It would be desirable to have means to better control the concentration of the protein under study (see discussion below). Alternatively, cells can be grown

on an unlabeled medium, harvested and then re-suspended in labeled media just before induction. The best in-cell NMR spectra are obtained with  $^{15}\text{N}$  labeled and deuterated rich media.  $^{13}\text{C}$  labeled proteins have been used much less because of the strong background resonances which make identification of useful resonances problematic [27]. Considerable improvement in the identification of the  $^{13}\text{C}$  resonances of the macromolecule has been achieved by selective labeling procedures, e.g. by adding methionine  $^{13}\text{C}$  labeled in the methyl group or leucine  $^{13}\text{C}$  labeled in the  $\delta$ -methyl groups [27].

#### *The risk of leakage*

Some of the early experiments of in-cell NMR suffered from the interference of leakage. Very often the NMR spectra of labeled proteins overexpressed in *E. coli* show broad resonances, even beyond detectability [28]. While this effect is certainly undesired, to see a sharp in-cell spectrum can be an alarm bell. Bacterial cells have a limited lifespan and burst open with aging. When this happens, the overexpressed protein is released into the cell suspension medium yielding spectra indistinguishable from those *in vitro*. As described by Barnes and Pielak [29] this problem can be avoided by careful controls and by limiting the concentration of the protein and the time of the experiment. They demonstrated that leakage becomes likely when the amount of protein is close to or larger than 50 fg/cell. Experiments should be also performed within very few hours from induction. It is anyway always necessary to prove that the cells have remained intact by spinning down the cells and verifying the spectrum of the supernatant. Spillage and its amount can be easily detected in this way.

#### *Expressing proteins in Eukaryotes*

In-cell NMR in eukaryotic cells is an entirely different story. As a rule, it is more problematic to overexpress labeled proteins in eukaryotic cells, but the different characteristics of eukaryotic cells, mainly their sheer larger dimensions and the greater resistance to leakage, offered new possibilities for studying proteins at reasonable concentrations in a natural environment. Eukaryotic cells are also more meaningful when the protein under study is a eukaryotic one.

The simplest way to introduce labeled proteins inside eukaryotic cells is to inject the protein solution in individual cells. Direct injection of a protein into cells has two main advantages with respect to over-expression: the concentration of the protein can be controlled *a priori* and the background generated by over-expression, arising from unwanted labeled metabolites, is effectively eliminated. The main disadvantage is that such a procedure, introduced by Selenko *et al.* [30] is possible only with very large eukaryotic cells like *Xenopus laevis* oocytes, but there are also other

non-trivial hurdles. The most critical step is the preparation and storage of very concentrated solutions: it is essential to use very concentrated solutions to insert proteins in oocytes because the injection causes a large dilution (approximately 20-fold). Besides, the procedure is cumbersome and not precisely reproducible because of the inherent variability in the selection of oocytes.

A second method of potentially general applicability is to use cell-penetrating peptides (CPPs) as a “Trojan Horse” to smuggle the protein inside the cell [31, 32]. The labeled protein is conjugated with a peptide that can puncture the cell membrane and thus enter the cell. CPPs are short peptides of sequences comprised between 10 and 30 amino acids, generally rich in basic residues which were already in use to shuttle inside cells various cargos, such as proteins, peptides, oligonucleotides before the advent of in-cell NMR. Inomata *et al.* [33] were the first to apply this methodology to in-cell NMR. They fused a cell-penetrating peptide tag onto three recombinant labeled proteins to transduce them into HeLa and COS-7 cells. After this ‘cargo’ delivery and removal of the peptide, the authors were able to record very high-quality NMR spectra inside mammalian cells. The advantage of this method is again that it is possible to control the concentration of the protein, but the method is complex and laborious, also because it would be preferable to cut the peptide from the cargo protein after cell penetration.

Alternatively, it is possible to make cells permeable to proteins using pore-forming bacterial toxins and thus enable spontaneous infusion of isotope-labeled proteins [31, 34]. The main advantage of this approach is that it does not require complex and costly modifications of the protein. The main disadvantage is that it can only be applied to cells grown in suspension and on very soluble isotope-labeled proteins. It must also be borne in mind that exposing mammalian cells to pore-forming toxins can be harmful.

Recently, yet another general method was proposed by Beata Bekei who, in her Dissertation thesis [35], describes the use of electroporation to introduce proteins in cells. To understand the molecular mechanism of electroporation it is convenient to subdivide it into several steps. The first step, also called induction, is the application of a pulse of an extracellular electric field. The subsequent maintenance of an overcritical electrical field is known as the expansion step. The step during which the electric field decreases is known as a stabilization step. The two final steps involve membrane resealing and the so-called memory effect. This effect describes the fact that, although most cells show normal behavior after the electroporation procedure, it is possible that certain changes in the intracellular cytoskeleton network are permanent. This method presents two main advantages with respect to peptide- and toxin-mediated protein delivery: it does not require any chemical modification of the protein under study and, even more important, it works without treating the cells with potentially harmful agents that can impair normal cell functioning. In addition

to the introduction to electroporation, the thesis of Beata Bekei [35] gives very useful full details of all previous techniques to insert proteins in eukaryotic cells.

### **How crowding affects in-cell NMR spectra**

It is now clear that the number of folded proteins that, when overexpressed in bacteria, show a good quality NMR spectrum is scanty [13]. There are, in principle, many possible causes for this partial failure, but the culprit pointed at most frequently is crowding. How can crowding influence in-cell NMR? The immediate cause of the difficulty of observing NMR signals of properly folded proteins in intact bacterial cells can be ascribed to the mobility of the protein molecule, particularly rotational freedom. Macromolecular crowding can affect mobility either directly, by influencing the viscosity of the solution or indirectly by stabilizing folded species with respect to unfolded species. Let us examine these two aspects systematically, with the aid of the (many) literature data now available.

In its simplest and qualitative form, theory predicts that neutral macromolecular crowders, i.e. globular macromolecules that do not interact with the protein examined, should increase the population of the folded species over that of the unfolded species because the latter is supposed to occupy a larger volume [1, 36]. During the last few years it has become increasingly clear that most, if not all, synthetic polymers generally employed as crowders are not really neutral [37-39]. It is imperative that neutrality of any crowder with respect to the macromolecule under study ought to be checked experimentally before undergoing any structural study in a crowded environment. At the same time Pielak and coworkers have drawn attention to the role played by soft interactions when (more realistic) protein crowders are employed instead of synthetic polymers [40, 41]. Weak non-specific forces can either stabilize or destabilize proteins. There are even more basic considerations that tend to downsize the indirect influence of crowding. It has been demonstrated that volume changes in the high temperature transition between folded and unfolded species occurring at ordinary atmospheric pressure can be much smaller than previously thought and, even more important, can lead not only to an increase but also to a decrease of volume [42]. When it has been possible to compare low and high temperature unfolded species, it was shown that the low temperature one is more expanded than the corresponding high temperature species [43-45]. This is even more dramatic when transitions are induced by elevated pressures [46]. Pressures of the order of thousands of atmospheres destabilize folded proteins both at high and low temperature, leading to species of smaller volume, albeit less *compact* than the folded one. The direct influence of viscosity has been thoroughly examined by Wang et al. [47]. These authors used C12 as test protein and made very careful comparisons of the influence of protein and synthetic crowders. At the higher

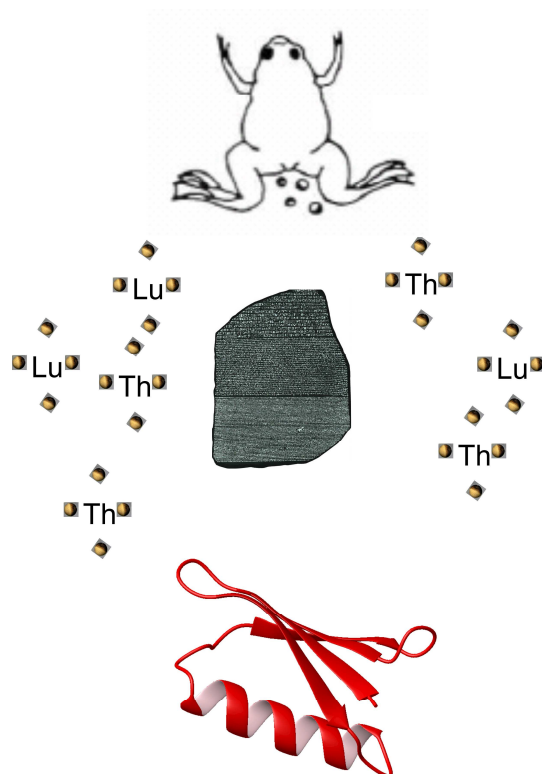
concentration used (300g/L), lysozyme, ovalbumin, and lysates make the HSQC spectrum of C12 disappear. Perhaps, the most surprising result of the systematic investigation by Wang et al. [47] is that the increase in viscosity caused by synthetic polymers affects diffusion of C12 in an intrinsically different way. Synthetic crowders slow down the translational motion more than its rotational motion whereas proteins have the opposite influence. The study by Wang et al. [47] showed convincingly that weak forces among proteins can play a dominant role in making in-cell NMR spectra generally difficult to observe. An essential *caveat* is the exceedingly high concentration of the crowder proteins employed (300g/L corresponding to ca. 21 mM in the case of lysozyme). At this concentration, most proteins tend to strongly *self-aggregate*. For instance, Price et al. [48] estimated that the equilibrium constant for the self-association of lysozyme at pH 4.6 and 298 K in the presence of 0.5 M NaCl is  $118 \text{ M}^{-1}$ . This constant tells us that at least half of the lysozyme molecules are aggregated for a concentration of the order of 21 mM. Thus, it is not surprising that the protein is incorporated into aggregates of the crowder and unable to rotate. The use of an un-physiological concentration of a *single* protein does not diminish the possible importance of weak non-specific interactions. In addition, it has been shown that cytosolic proteins are intrinsically sticky [49]. However, it is more likely that in an environment rich of several different proteins the drastic decrease of rotational tumbling is caused by a combination of weak protein interactions and the presence of very rigid macromolecules like those of nucleic acids [50] or of a protein complex [51].

### **In-cell structure determination**

Given the discussed limitations, it is not surprising that there are not many structure determinations of proteins in-cell. The only complete structure determination by in-cell NMR is that published by Sakakibara et al. [12]. These authors solved the structure of TTHA1718, a putative heavy-metal binding protein from *Thermus thermophilus* HB8 by in-cell NMR. This feat required the overcoming of considerable technical hurdles. Most of all, to overcome problems originating from the instability of living cells and the intrinsic low sensitivity of in-cell experiments, the authors had to drastically reduce measurement time in 3D NMR spectra by nonlinear sampling of the indirectly acquired dimension. They were able to assign almost all the backbone and most of the side-chain atom NMR resonances, thus calculating high quality structures of TTHA1718 which are very similar to the previously determined in vitro structure. The difficulties encountered and overcome by Sakakibara et al. [12] possibly explain why no other full structure has been determined since by in-cell NMR. A second “historical” case include the complete NMR assignment of GB1, achieved via the in vivo implementation of a suite of fast 3D NMR experiments [52]. Curiously, this article

described the preliminary step in NMR structure determination, but was not followed by an explicit structure determination.

Recently, a new interesting approach to structure determination by in-cell NMR was described. Instead of using the traditional sequence of 3D NMR experiments generally employed in *in vitro* determinations, Müntener et al. [53] used a combination of 2D experiments on a protein tagged with a paramagnetic probe.



**Figure .1** Avery high resolution structure of GB1 in *Xenopus* oocytes was determined interpreting RDC and PCS data with Rosetta software. DOTA tagged samples injected in oocytes contained different lanthanides (Lu and Th in the figure).

They introduced three modified tetraaza-carboxylic (DOTA) chelators into the GB1 domain (GB1) to bind either diamagnetic Luthetium (Lu) or paramagnetic Thulium (Tm). When these modified protein samples were hosted in intact *Xenopus laevis* oocytes it was possible to simultaneously measure pseudocontact shifts (PCSs) and residual dipolar couplings (RDCs) (Figure 1). When used as input for structure calculation routines within the Rosetta program the parameters measured from a single set of 2D in-cell NMR experiments led to well-defined GB1 ensembles. The in-cell structures calculated via the Rosetta software proved closer to the X-ray structure (pdb id: 2QMT) than high resolution *in vitro* NMR structures (2PLP). In addition, it is worth noting that Müntener et al. [53] could use in-cell NMR samples of low concentration ( $\sim 50 \mu\text{M}$ ) and a moderate magnetic field strength (600 MHz) accessible to many laboratories.

In the wake of the paper by Muentner et al [53], several similar determinations closely followed [54-56].

### Folding and unfolding

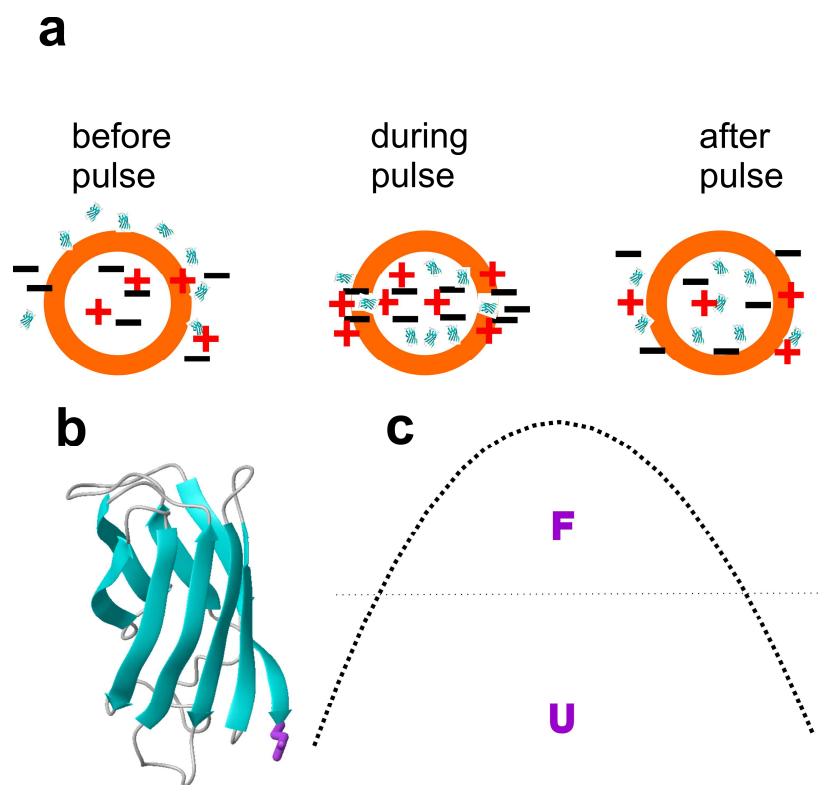
A very important target of in-cell studies is the assessment of protein stability *in vivo*. The response of in-cell NMR studies is mixed, as much as that of corresponding studies in crowded solutions, varying from no effect at all to a destabilizing influence. Before examining some of the in-cell NMR works in detail, it may be useful to set the pace by quoting a very accurate in-cell study performed by Koenig et al. [57] with a single-molecule FRET technique. These authors took advantage of the possibility, offered by the yeast analog of frataxin, to measure the *whole stability curve* of the protein [43, 58, 59]. They found that the data in HeLa cells reflected the behavior observed *in vitro* [44, 58], in spite of the lower signal-to-noise ratio. In other words, stability was apparently not changed significantly by the cell environment.

Schlesinger et al. [60] tested the widespread belief that volume exclusion dominates the crowding effect in cells using a strongly destabilized mutant of protein L. Seven lysine residues present in the wild type protein were replaced by glutamic acids, thus increasing the population of unfolded protein in dilute solution from 0.1 to 84%. This mutated construct folds reversibly upon addition of Na<sup>+</sup> or K<sup>+</sup> ions. Using in-cell NMR spectroscopy the authors showed that the cytoplasm of *E. coli* does not overcome even the modest (~1 kcal/mol) free-energy deficit corresponding to the population change. This experiment certainly proved that excluded-volume effects alone are not sufficient to reverse the population change of the mutated protein L construct. However, the direct comparison of salt neutralization of the seven glutamic residues and volume-exclusion effects seems a bit unfair because ionic strength changes and volume exclusion are hardly comparable.

In-cell thermodynamics, a trendy topic, may be considered an emerging subfield in protein stability by in-cell NMR. The two most relevant papers in this field have been published by the groups of Pielak and Oliveberg. Smith et al. [61] used fluorine NMR data on the 7-kDa globular N-terminal SH3 domain of *Drosophila* signal transduction protein drk (SH3) and found that the cell environment of *E. coli* can modulate stability, leading to a decrease, an increase or no effect whatsoever depending on properties of the protein surface.

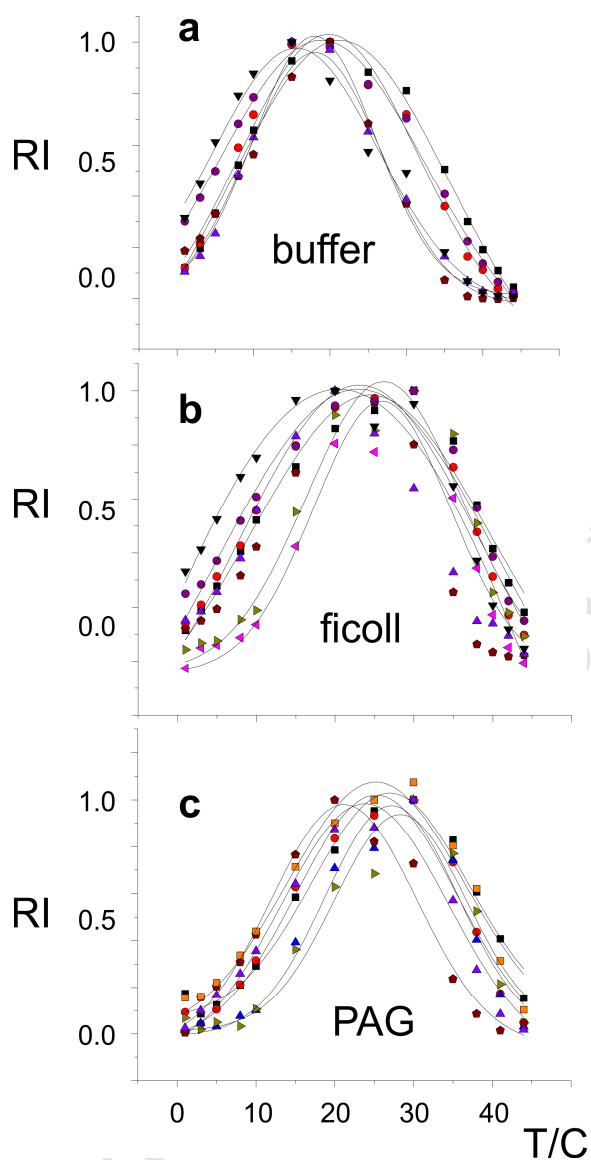
As shown by interactions with several crowders, charge–charge interactions are fundamental to protein stability and folding kinetics in cells. Danielsson et al. [62] studied a mutant of the so-called SOD1<sup>barrel</sup>, a variant of the ubiquitous radical scavenger Cu/Zn superoxide dismutase. The I35A mutant of SOD1<sup>barrel</sup>, SOD1<sup>I35A</sup> is so strongly destabilized that, at room temperature *in vitro*, exists as a mixture of folded and unfolded species, an ideal condition to measure both cold and heat

denaturation, thus allowing the measurement of the whole stability curve. Danielsson et al. [62] found that SOD1<sup>I35A</sup> is further destabilized both in *E.coli* and in cultured A2780 cells (Figure 2).



**Figure 2.** In-cell thermodynamics. A strongly destabilized construct of super oxide dismutase, SOD<sup>I35A</sup>, was introduced in cultured A2780 cells by means of electroporation. Plus and minus signs refer to electric charges. (a). The NMR HSQC spectrum shows two peaks of the C terminal Q110 (b, violet side chain) corresponding to the folded and unfolded species. Measurements of NH cross peak volumes allowed accurate evaluation of the stability curve in several conditions (c).

The main conclusion was that it is not possible to predict protein stability in cells on the sole basis of its sequence because it is necessary to take into account the specific intracellular environment. This brilliant work paves the way for future in-cell NMR studies. The only *caveat* is that all experimental findings were based on the volume changes of a single cross peak of SOD1<sup>I35A</sup>, i.e. that of the C terminal Q110 NH, as measured in HSQC spectra. It is well known that the measurement of peak volumes in 2D NMR spectra is plagued with difficulties [63] and is particularly influenced by the medium viscosity. It is thus possible that the different viscosity and their temperature dependence played an unaccounted role. In addition, residues far from the protein core may give contradictory values of protein stability. For instance, when trying to compare the influence of crowding and confinement on the stability of Yfh1 using the volumes of peaks from HSQC experiments we observed wild variations of relative peak intensities (RI) as a function of temperature when taken from different residues (Figure 3).



**Figure 3.** Relative intensities (RI) of several cross peaks of Yfh1 in three different environments: buffer (a), 15% Ficoll 70 (b) and 8% PAG (c).

It is clear why volume measurements of HSQC cross peaks could not be used to calculate protein stability curves in this case [59]. It is in order to note that, although the above and several other articles stress the importance of quinary interactions, they cannot be considered *ipso facto* has a proof of a negligible effect of volume exclusion. It is conceivable that in cell we observe the balance of numerous, often contrasting influences.

Another field in which in-cell NMR can play a decisive role is the conformational stability of intrinsically unfolded proteins (IDPs). The main aspects of in-cell NMR of IDPs are treated in great detail in a recent review by Theillet et al. [21] but it is still worth repeating here the main aspects. IDPs constitute a special case for in-cell NMR, mainly because of their favorable dynamic properties that lead to superior in-cell NMR qualities [13]. This aspect has been exploited to resolve complicated issues on the true state of  $\alpha$ -synuclein, a protein widely studied by the NMR

community. For many years  $\alpha$ -synuclein has been considered a typical IDP, both *in vitro* and *in vivo*, but this vision was challenged by the groups of Selkoe and Petsko who argued that the true tetrameric nature of the protein *in vivo* is destroyed during the purification steps to produce the recombinant protein [64, 65]. This hypothesis ignored previous *in-cell* NMR reports on  $\alpha$ -synuclein which had demonstrated that it is monomeric and unfolded in living *E. coli* cells [28, 66, 67]. Recent *in-cell* NMR studies [68] confirmed, in agreement with earlier studies, that  $\alpha$ -synuclein is intrinsically disordered and monomeric not only *in vitro* but also in live bacteria.

Selective visibility in *in-cell* NMR of intrinsically disordered segments of an otherwise folded protein can be exploited to investigate the state of a protein in cell (aggregate or monomeric) and clarify the role of an import signal. Popovic et al. [69] have employed *in-cell* NMR to compare the behavior of orthologs of the frataxin family, ubiquitously present in prokaryotes and in eukaryotes. All orthologs contain a folded domain, which in eukaryotes is preceded by an N-terminal peptide acting as the mitochondrial import signal. They showed that the HSQC NMR spectrum of the bacterial ortholog CyaY is not visible in *E. coli* cells, but becomes fully observable as soon as the cells are lysed. On the contrary, the NMR spectrum of the yeast ortholog Yfh1 contains visible peaks from the protein. These peaks correspond to the flexible N-terminal peptide, proving that it is flexible and disordered. The flexibility of the N-terminal peptide is consistent with previous studies of human frataxin, despite the sequence diversity of this region in the two proteins. In addition, the results of Popovic et al. [69] show that, in cell, the protein does not exist as an aggregate but as a monomeric species.

Also in the case of IDPs it is possible to find controversial views on the influence of the cell environment on stability and folding. For instance, Dedmon et al. [70] claimed that FlgM, an intrinsically unstructured protein becomes structured in cell, but their evidence is essentially based on negative data: the protein peaks in the HSQC spectrum of FlgM in cell are not observable [41]. Free FlgM a protein from *Salmonella typhimurium*, which regulates flagellar synthesis appears unstructured in buffer solution, but its C-terminal half can form a helix upon binding to the transcription factor  $\Delta 28$  as hinted at by the disappearance of some C-terminal NMR peaks. In their study, Dedmon et al. [70] found that also in *E. coli*, the same NMR peaks disappear. In their interpretation, this is evidence that the cellular environment modulates the structural *in vivo* properties of this disorderd protein, although other interactions might similarly explain the disappearance of FlgM NMR signals.

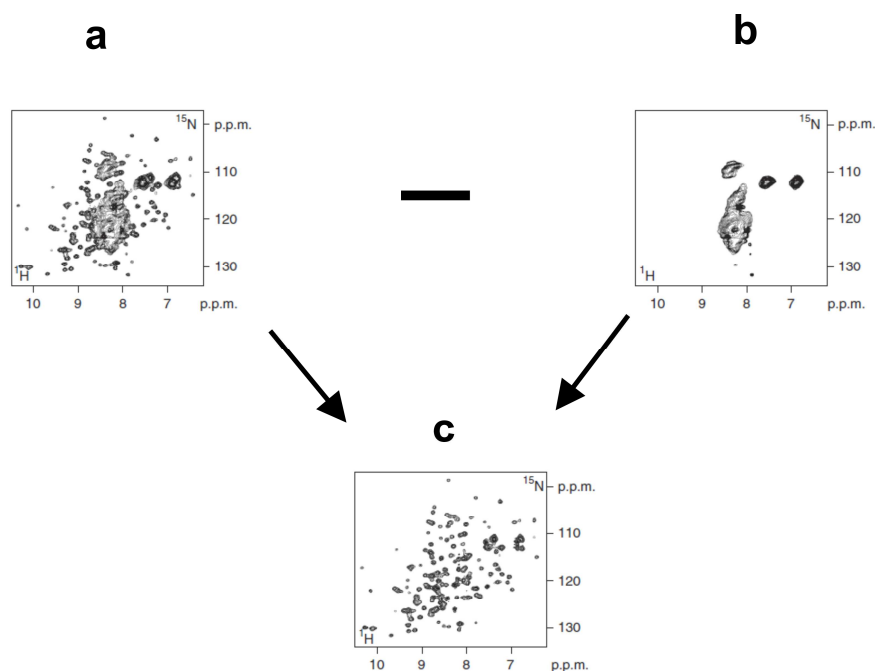
## Functional aspects

A very promising application of in-cell NMR is the *in vivo* study of interactions among metabolites of all kinds. Among the few NMR methods devised specifically for in-cell NMR stand out a series of techniques aimed at recording interactions *in vivo*, STINT [71] and the more recent SMILI [18]. These methods are not new spectroscopic methods but rather procedures to focus on the sole interactions of interest. To map the structural interactions between protein partners which lead to complex formation, Burz *et al.* [71] developed an in-cell NMR method dubbed STINT-NMR. The method is based on the possibility of expressing the proteins of the complex in a time-controlled manner [72] instead of expressing them simultaneously. As a consequence, recording their NMR spectra as a function of time yields a titration of the interaction, suggesting crucial details of the interacting sides. The time course of the experiment is as follows: first a target protein is overexpressed in  $U\text{-}^{15}\text{N}$  medium, then the growth medium is changed to an unlabeled one and the interactor protein is expressed. If the structure and NMR assignment of the target protein are known, it is comparatively easy to interpret the changes in line widths and chemical shifts of peaks of the target protein in terms of facing surfaces in the complex. Burz *et al.* [71] illustrated this method with  $^{15}\text{N}$  ubiquitin and two ubiquitin ligands: a 28-amino acid peptide from ataxin 3 (AUIM) and the Signal-Transducing Adapter Molecule (STAM2) [73] expressed in unlabeled medium. The two main factors limiting the use of STINT are the concentration of the target protein and the difficulty to control the integrity of the interacting proteins. Altogether, STINT looked as a very promising method, but there has been no direct follow up since 2006, possibly because it is so difficult to find proteins that can be seen by in-cell NMR. Another cause of concern about this method is that, in the paper describing the method [71], the authors claim that freeze-thawing the samples for a month does not cause cell lysis. Such an observation contradicts the experience of most laboratories working on in-cell NMR. The same group, shortly after proposing STINT, developed a method to screen small molecules directly in cell Xie *et al.* [74]: it was called Screening of small Molecule Interactor Library (SMILI-NMR). SMILI-NMR utilizes the STINT NMR [66] technology to produce complexes inside the cell with one protein uniformly  $^{15}\text{N}$  labeled. Changes in NMR spectra caused by binding of small potential ligands are analyzed in terms of their relationship with the stability of the complex and can reveal biologically relevant, functional interaction surfaces. SMILI-NMR provides an important means to bridge the gap between biochemical identification of small ligands capable of interfering with complexes and the biological activity resulting from the inhibition of cellular processes by these ligands. The method requires minimal sample preparation and eliminates the need for extensive protein purification. Furthermore, SMILI-NMR can be automated by making use of robotic high throughput screening (HTS)

accessories available for modern NMR spectrometers, such as liquid handlers and NMR tube changers. There have not been many follow-up STINT experiments in prokaryotic cells. The most prominent recent application is still from some of the original proponents [75].

In recent years, efficient intracellular protein expression, once possible mainly in prokaryotic cells, has been extended to mammalian cells [76, 77], yeast [78] and insect cells [79]. One of the most exciting aspects of these improved expression techniques is that, at least in principle, it is possible to express proteins localized in different cellular compartments, and thus investigate the effect of different subcellular environments [78]. The possibility of intracellular protein expression was exploited by Banci and coworkers to study relevant functional aspects such as protein maturation or redox-controlled protein fold in living cells. Altogether this group has produced an impressive number of papers [80-86]. The main aspects of these works are summarized in great detail in a recent paper [87].

The most important application described in this methodological paper is the comparative study of superoxide dismutase 1 (SOD1) and a few of its mutants linked to familial amyotrophic lateral sclerosis (fALS). Cu/Zn-superoxide dismutase (SOD1) is a radical scavenger that may misfold and then aggregate in the neurons of people affected by ALS. It all started with a study of SOD1 maturation, a process that consists of several steps: zinc binding, dimerization, inclusion of copper and oxidation of the disulfide bond that holds the two subunits. Intracellular protein expression allowed the observation of different protein states, from the apo species to various metal-containing and oxidized forms leading to the final mature protein. Addition of Zn ions to the culture medium eliminates apo-SOD1, the only species present, leading to a dimeric species which hosts one Zn ion per protomer. The spectra of this species (Figure 4) have very good resolution, albeit obtained with a rather cumbersome procedure. The  $^{15}\text{N}$ -labeled cell samples gave rise to strong background signals, due to the presence of other labeled proteins and peptides, notably glutathione. These background signals had to be reduced by the subtracting two transformed spectra: one acquired on the in-cell NMR sample containing the protein of interest, and the other acquired on a control in-cell NMR sample, prepared in exactly the same conditions as the previous one but with the cells transfected with the empty DNA vector. Spectra subtraction is a common procedure in NMR but, usually, it is performed on the same sample tube after altering some parameter inside the tube. The procedure adopted by Barbieri et al. [87], in principle, is very tricky because it involves two different sample tubes and it may be difficult to choose exactly identical conditions. However, their procedure is so successful that the final difference spectrum has a resolution comparable to that of in vitro spectra of the same protein [87] (Figure 4).



**Figure 4.** Spectra subtraction for background elimination. a) SOFAST HMQC spectrum of SOD1 in cells. b) SOFAST HMQC spectrum of a control cell sample with an empty DNA vector. c) subtraction of a-b. Elaborated from reference [87]. {necessary to ask permission}

As it was well known, copper incorporation could not proceed without the presence of another protein called CCS. Barbieri et al. [87] observed that *in vivo* the role of CCS is essential for both copper incorporation and disulfide bond formation. Thus, it proved possible to probe the behaviour of several SOD1 mutants linked to familial ALS and, eventually, the accumulation of unstructured species in the cytoplasm. The maturation levels of fALS mutants in different cellular conditions could be compared with the corresponding ones of hSOD1 and then analysed in terms of their ability to remain in solution without forming aggregates. They found that in some fALS-linked mutants Zn-SOD1 was not present or had low concentration suggesting that the mutations can have a negative influence on SOD1 maturation. In addition, many of the fALS mutants examined were present as unfolded species, both in-cell and in vitro. However, the same mutants were correctly folded if the copper chaperon (hCCS) was co-expressed in the presence of copper and zinc ions. This finding was interpreted as a proof that unfolded species do not form when the maturation mechanism is operative. Luchinat et al. [83] proposed that the unfolded species is generated from the monomeric apo species, which in turn gives rise to aggregates typical of fALS, supporting a mechanism for the onset of this disease in which, some mutations lead to an increased population of unstructured apo form, which acts can give rise to toxic oligomeric species.

## Conclusions

We have discussed the importance and limitations of a technique that in principle could provide unique information about the state of folding, structure and stability of proteins directly in their milieu. It is clear that, despite of the potentialities of the technique, the number of proteins which, at the moment, allow NMR spectra visible in living prokaryotic cells is limited. In-cell NMR in eukaryotic cells is more promising and likely constitutes the frontiers of the technique even though it confronts us with important but interesting challenges. Much more work is needed to open new potentialities that may enable us to follow proteins in their natural environment.

## References

- [1] H. X. Zhou, G. Rivas, A. P. Minton, Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences, *Annu. Rev. Biophys.* 37 (2008) 375-97.
- [2] S. B. Zimmerman, S. O. Trach, Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*, *J. Mol. Biol.* 222 (1991) 599-620.
- [3] K. Luby-Phelps, Cytoarchitecture and physical properties of cytoplasm, volume, viscosity, diffusion, intracellular surface area. *Int. Rev. Cytol.* 192 (2000) 189-221.
- [4] R. J. Ellis, Macromolecular crowding: an important but neglected aspect of the intracellular environment *Curr. Opin. Struct. Biol.* 11 (2001) 114-9.
- [5] H. X. Zhou, Protein folding and binding in confined spaces and in crowded solutions, *J. Mol. Recognit.* 17 (2004) 368-75.
- [6] G. Rivas, A. P. Minton, Macromolecular Crowding In Vitro, In Vivo, and In Between, *Trends Biochem. Sci.* 41 (2016) 970-981.
- [7] A. Politou, P. A. Temussi, Revisiting a dogma: the effect of volume exclusion in molecular crowding, *Curr. Opin. Struct. Biol.* 30 (2015) 1-6.
- [8] P. S. Allen, R. B. Thompson, A. H. Wilman, Metabolite-specific NMR spectroscopy in vivo. *NMR Biomed.* 10 (1997) 435-44.
- [9] Z. Serber, L. Corsini, F. Durst, V. Dötsch, In-cell NMR spectroscopy, *Methods Enzymol.* 394 (2005) 17-41.
- [10] Z. Serber, R. Ledwidge, S. M. Miller, V. Dötsch, Evaluation of parameters critical to observing proteins inside living *Escherichia coli* by in-cell NMR spectroscopy, *J. Am. Chem. Soc.* 123 (2001) 8895-901.

- [11] K. E. Robinson, P. N. Reardon, L. D. Spicer, In-cell NMR spectroscopy in *Escherichia coli*, *Methods Mol. Biol.* 831 (2012) 261-77.
- [12] D. Sakakibara, A. Sasaki, T. Ikeya, J. Hamatsu, T. Hanashima, M. Mishima, M. Yoshimasu, N. Hayashi, T. Mikawa, M. Wälchli, B. O. Smith, M. Shirakawa, P. Güntert, Y. Ito, Protein structure determination in living cells by in-cell NMR spectroscopy, *Nature* 458 (2009) 102-5.
- [13] H. Tochio, Watching protein structure at work in living cells using NMR spectroscopy, *Curr. Opin. Chem. Biol.* 16 (2012) 609-13.
- [14] Z. Serber, V. Dötsch, In-cell NMR spectroscopy. *Biochemistry.* 40 (2001) 14317-23.
- [15] S. Reckel, F. Löhr, V. Dötsch, In-cell NMR spectroscopy, *Chembiochem.* 6 (2005) 1601-6.
- [16] P. Selenko, G. Wagner, Looking into live cells with in-cell NMR spectroscopy, *J. Struct. Biol.* 158 (2007) 244-53.
- [17] V. Dötsch, Investigation of proteins in living bacteria with in-cell NMR experiments, *Top Curr Chem.* 273 (2008) 203-14.
- [18] A. Y. Maldonado, D. S. Burz, A. Shekhtman, In-cell NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* 59 (2011) 197-212.
- [19] C. Li, M. Liu, Protein dynamics in living cells studied by in-cell NMR spectroscopy, *FEBS Lett.* 587 (2013) 1008-11.
- [20] R. Hänsel, L. M. Luh, I. Corbeski, L. Trantirek, V. Dötsch, In-cell NMR and EPR spectroscopy of biomacromolecules, *Angew. Chem. Int. Ed. Engl.* 53 (2014) 10300-14.
- [21] F. X. Theillet, A. Binolfi, T. Frembgen-Kesner, K. Hingorani, M. Sarkar, C. Kyne, C. Li, P. B. Crowley, L. Gierasch, G. J. Pielak, A. H. Elcock, A. Gershenson, P. Selenko, Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs), *Chem. Rev.* 114 (2014) 6661-714.
- [22] D. I. Freedberg, P. Selenko, Live cell NMR, *Annu. Rev. Biophys.* 43 (2014) 171-92.
- [23] M. J. Smith, C. B. Marshall, F. X. Theillet, A. Binolfi, P. Selenko, M. Ikura, Real-time NMR monitoring of biological activities in complex physiological environments, *Curr. Opin. Struct. Biol.* 32 (2015) 39-47.
- [24] E. Luchinat, L. Banci, A Unique Tool for Cellular Structural Biology: In-cell NMR, *J Biol Chem.* 291 (2016) 3776-84.
- [25] C. Li, G. F. Wang, Y. Wang Y, R. Creager-Allen, E. A. Lutz, H. Scronce, K. M. Slade, R. A. Ruf, R. A. Mehl, G. J. Pielak, Protein (19)F NMR in *Escherichia coli*, *J. Am. Chem. Soc.* 132 (2010) 321-7.
- [26] P. B. Crowley, C. Kyne, W. B. Monteith, Simple and inexpensive incorporation of 19F-tryptophan for protein NMR spectroscopy, *Chem. Commun. (Camb)* 48 (2012) 10681-3.

- [27] Z. Serber, W. Straub, L. Corsini, A. M. Nomura, N. Shimba, C. S. Craik, P. Ortiz de Montellano, V. Dötsch, Methyl groups as probes for proteins and complexes in in-cell NMR experiments, *J. Am. Chem. Soc.* 126 (2004) 7119-25.
- [28] C. Li, L. M. Charlton, A. Lakkavaram, C. Seagle, G. Wang, G. B. Young, J. M. MacDonald, G. J. Pielak, Differential dynamical effects of macromolecular crowding on an intrinsically disordered protein and a globular protein: implications for in-cell NMR spectroscopy, *J. Am. Chem. Soc.* 130 (2008) 6310–6311.
- [29] C. O. Barnes, G. J. Pielak, In-cell protein NMR and protein leakage, *Proteins* 79 (2011) 347-51.
- [30] P. Selenko, Z. Serber, B. Gadea, J. Ruderman, G. Wagner, 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 (2006) 11904.
- [31] S. Ogino, S. Kubo, R. Umemoto, S. Huang, N. Nishida, I. Shimada, 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 (2009) 10834.
- [32] J. Danielsson, K. Inomata, S. Murayama, H. Tochio, L. Lang, M. Shirakawa, M. Oliveberg, Pruning the ALS-associated protein SOD1 for in-cell NMR, *J. Am. Chem. Soc.* 135 (2013) 10266.
- [33] K. Inomata, A. Ohno, H. Tochio, S. Isogai, T. Tenno, I. Nakase, T. Takeuchi, S. Futaki, Y. Ito, H. Hiroaki, M. Shirakawa, High-resolution multidimensional NMR spectroscopy of proteins in human cells, *Nature* 458 (2009) 106-109.
- [34] S. Kubo, N. Nishida, Y. Udagawa, O. Takarada, S. Ogino, I. Shimada, A gel-encapsulated bioreactor system for NMR studies of protein-protein interactions in living mammalian cells, *Angew. Chem., Int. Ed.* 52 (2013) 1208-11.
- [35] B. Bekei, *In-cell NMR Spectroscopy in Mammalian Cells*. Ph.D. thesis, Freie Universität Berlin, Germany, (2013)
- [36] A. P. Minton, Effect of a concentrated "inert" macromolecular cosolute on the stability of a globular protein with respect to denaturation by heat and by chaotropes: a statistical-thermodynamic model, *Biophys. J.* 78 (2000) 101-9.

- [37] P. B. Crowley, K. Brett, J. Muldoon, NMR spectroscopy reveals cytochrome c-poly(ethylene glycol) interactions, *ChemBioChem*. 9 (2008) 685–688.
- [38] M. Jiao, H. T. Li, J. Chen, A. P. Minton, Y. Liang, Attractive protein-polymer interactions markedly alter the effect of macromolecular crowding on protein association equilibria. *Biophys. J.* 99 (2010) 914-23.
- [39] J. M. Paredes, L. Crovetto, A. Orte, S. G. Lopez, E. M. Talavera, J. M. Alvarez-Pez, Photophysics of the interaction between a fluorescein derivative and Ficoll, *J. Phys. Chem. A*. 115 (2011) 13242-50.
- [40] Y. Wang, M. Sarkar, A. E. Smith, A. S. Krois, G. J. Pielak, Macromolecular crowding and protein stability, *J. Am. Chem. Soc.* 134 (2012) 16614-8.
- [41] A. E. Smith, Z. Zhang, G. J. Pielak, C. Li, NMR studies of protein folding and binding in cells and cell-like environments, *Curr. Opin. Struct. Biol.* 30 (2015) 7-16.
- [42] T. Chalikian, R. Filfil, How large are the volume changes accompanying protein transitions and binding? *Biophys. Chemistry*, 104 (2003) 489–499.
- [43] M. Adrover, V. Esposito, G. Martorell, A. Pastore, P. A. Temussi, Understanding cold denaturation: the case study of Yfh1, *J. Am. Chem. Soc.* 132, (2010) 16240-6.
- [44] M. Aznauryan, D. Nettels, A. Holla, H. Hofmann, B. Schuler, Single-molecule spectroscopy of cold denaturation and the temperature-induced collapse of unfolded proteins, *J. Am. Chem. Soc.* 135 (2013) 14040-3.
- [45] B. Vilanova, D. Sanfelice, G. Martorell, P. A. Temussi, A. Pastore, Trapping a salt-dependent unfolding intermediate of the marginally stable protein Yfh1, *Front. Mol. Biosci.* 1 (2014) 13.
- [46] A. Zipp, W. Kauzmann, Pressure denaturation of metmyoglobin, *Biochemistry* 12 (1973) 4217.
- [47] Y. Wang, C. Li, G. J. Pielak, Effects of proteins on protein diffusion, *J. Am. Chem. Soc.* 132 (2010) 9392-7.
- [48] W. S. Price, F. Tsuchiya, Y. Arata, Time dependence of aggregation in crystallizing lysozyme solutions probed using NMR self-diffusion measurements, *Biophys. J.* 80 (2001) 1585-90.
- [49] P. B. Crowley, E. Chow, T. Papkovskaia, Protein interactions in the *Escherichia coli* cytosol: an impediment to in-cell NMR spectroscopy, *Chembiochem*. 12 (2011) 1043-8.
- [50] A. M. Augustus, P. N. Reardon, L. D. Spicer, MetJ repressor interactions with DNA probed by in-cell NMR, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 5065-9.

- [51] D. Sanfelice, M. Adrover, G. Martorell, A. Pastore, P. A. Temussi, Crowding versus molecular seeding: NMR studies of protein aggregation in hen egg white, *J. Phys. Condens. Matter* 24 (2012) 244107.
- [52] P. N. Reardon, L. D. Spicer, Multidimensional NMR spectroscopy for protein characterization and assignment inside cells, *J. Am. Chem. Soc.* 127 (2005) 10848-9.
- [53] T. Müntener, D. Häussinger, P. Selenko, F. X. Theillet, In-Cell Protein Structures from 2D NMR Experiments, *J. Phys. Chem. Lett.* 7 (2016) 2821-5.
- [54] B. B. Pan, F. Yang, Y. Ye, Q. Wu, C. Li, T. Huber, X. C. Su. 3D structure determination of a protein in living cells using paramagnetic NMR spectroscopy, *Chem. Commun. (Camb)*. 52(2016) 10237-40.
- [55] Y. Hikone, G. Hirai, M. Mishima, K. Inomata, T. Ikeya, S. Arai, M. Shirakawa, M. Sodeoka, Y. Ito. A new carbamidemethyl-linked lanthanoid chelating tag for PCS NMR spectroscopy of proteins in living HeLa cells, *J. Biomol. NMR.* 66 (2016) 99-110.
- [56] T. Ikeya, T. Hanashima, S. Hosoya, M. Shimazaki, S. Ikeda, M. Mishima, P. Güntert, Y. Ito. Improved in-cell structure determination of proteins at near-physiological concentration, *Sci. Rep.* 6 (2016) 38312.
- [57] I. König, A. Zarrine-Afsar, M. Aznauryan, A. Soranno, B. Wunderlich, F. Dingfelder, J. C. Stüber, A. Plückthun, D. Nettels, B. Schuler, Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells, *Nat. Methods* 12 (2015) 773-9.
- [58] A. Pastore, S. Martin, A. Politou, K. C. Kondapalli, T. Stemmler, P. A. Temussi, Unbiased cold denaturation: low and high temperature unfolding of Yeast frataxin under physiological conditions, *J. Am. Chem. Soc.* 129 (2007) 5374-5375.
- [59] D. Sanfelice, A. Politou, S. R. Martin, P. De Los Rios, P. A. Temussi, A. Pastore, The effect of crowding and confinement: A comparison of Yfh1 stability in different environments, *Phys Biol.* 10 (2013) 045002.
- [60] A. P. Schlesinger, Y. Wang, X. Tadeo, O. Millet, G. J. Pielak, Macromolecular crowding fails to fold a globular protein in cells, *J. Am. Chem. Soc.* 133 (2011) 8082-5.
- [61] A. E. Smith, L. Z. Zhou, A. H. Gorensek, M. Senske, G. J. Pielak, In-cell thermodynamics and a new role for protein surfaces, *Proc. Natl. Acad. Sci. U. S. A.* 113 (2016) 1725-30.
- [62] J. Danielsson, X. Mu, L. Lang, H. Wang, A. Binolfi, F. X. Theillet, B. Bekei, D. T. Logan, P. Selenko, H. Wennerström, M. Oliveberg, Thermodynamics of protein destabilization in live cells, *Proc. Natl. Acad. Sci. U S A* 112 (2015) 12402-7.

- [63][58] R. K. Rai, P. Tripathi, N. Sinha, Quantification of metabolites from two-dimensional nuclear magnetic resonance spectroscopy: application to human urine samples, *Anal. Chem.* 81 (2009) 10232-8.
- [64] T. Bartels, J. G. Choi, D. J. Selkoe,  $\alpha$ -Synuclein occurs physiologically as a helically folded tetramer that resists aggregation, *Nature* 477 (2011) 107-110.
- [65] W. Wang, I. Perovic, J. Chittuluru, A. Kaganovich, L. T. Nguyen, J. Liao, J. R. Auclair, D. Johnson, A. Landeru, A. K. Simorellis, S. Ju, M. R. Cookson, F. J. Asturias, J. N. Agar, B. N. Webb, C. Kang, D. Ringe, G. A. Petsko, T. C. Pochapsky, Q. Q. Hoang, A soluble  $\alpha$ -synuclein construct forms a dynamic tetramer, *Proc. Natl. Acad. Sci. U.S.A.* 108 (2011) 17797.
- [66] B. C. McNulty, G. B. Young, G. J. Pielak, Macromolecular crowding in the *Escherichia coli* periplasm maintains alpha-synuclein disorder, *J. Mol. Biol.* 355 (2006) 893-7.
- [67] R. L. Croke, C. O. Sallum, E. Watson, E. D. Watt, A. T. Alexandrescu, Hydrogen exchange of monomeric alpha-synuclein shows unfolded structure persists at physiological temperature and is independent of molecular crowding in *Escherichia coli*, *Protein Sci.* 17 (2008) 1434-45.
- [68] F. X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb, P. Selenko, Structural disorder of monomeric  $\alpha$ -synuclein persists in mammalian cells, *Nature* 530 (2016) 45-50.
- [69] M. Popovic, D. Sanfelice, C. Pastore, F. Prischi, P. A. Temussi, A. Pastore, Selective observation of the disordered import signal of a globular protein by in-cell NMR: the example of frataxins, *Protein Sci.* 24 (2015) 996-1003.
- [70] M. M. Dedmon, C. N. Patel, G. B. Young, G. J. Pielak, FlgM gains structure in living cells, *Proc. Natl. Acad. Sci. U. S. A.* 99 (2002) 12681-4.
- [71] D. S. Burz, K. Dutta, D. Cowburn, A. Shekhtman, In-cell NMR for protein-protein interactions (STINT-NMR), *Nat. Protoc.* 1 (2006) 146-52.
- [72] R. Lutz, H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements, *Nucleic Acids Res.* 25 (1997) 1203-1210.
- [73] K. G. Bache, C. Raiborg, A. Mehlum, H. Stenmark, STAM and HRS are subunits of a multivalent ubiquitin-binding complex on early endosomes, *J. Biol. Chem.* 278 (2003) 12513-12521.
- [74] J. Xie, R. Thapa, S. Reverdatto, D. S. Burz, A. Shekhtman, Screening of small molecule interactor library by using in-cell NMR spectroscopy (SMILI-NMR), *J. Med. Chem.* 52 (2009) 3516-3522.

- [75] A. Y. Maldonado, D. S. Burz, S. Reverdatto, A. Shekhtman, Fate of pup inside the *Mycobacterium* proteasome studied by in-cell NMR, *PLoS One*. 8 (2013) e74576.
- [76] A. R. Aricescu, W. Lu, E. Y. Jones, A time- and cost-efficient system for high-level protein production in mammalian cells, *Acta Crystallogr. D. Biol. Crystallogr.* 62 (2006) 1243–1250.
- [77] Y. Zhao, B. Bishop, J. E. Clay, W. Lu, M. Jones, S. Daenke, C. Siebold, D. I. Stuart, E. Y. Jones, A. R. Aricescu, Automation of large scale transient protein expression in mammalian cells, *J. Struct. Biol.* 175 (2011) 209–215.
- [78] K. Bertrand, S. Reverdatto, D. S. Burz, R. Zitomer, A. Shekhtman, Structure of proteins in eukaryotic compartments, *J. Am. Chem. Soc.* 134 (2012) 12798–806.
- [79] J. Hamatsu, D. O'Donovan, T. Tanaka, T. Shirai, Y. Hourai, T. Mikawa, T. Ikeya, M. Mishima, W. Boucher, B. O. Smith, E. D. Laue, M. Shirakawa, Y. Ito, High-resolution heteronuclear multidimensional NMR of proteins in living insect cells using a baculovirus protein expression system, *J. Am. Chem. Soc.* 135 (2013) 1688–1691.
- [80] L. Banci, L. Barbieri, I. Bertini, E. Luchinat, E. Secci, Y. Zhao, A. R. Aricescu, Atomic-resolution monitoring of protein maturation in live human cells by NMR, *Nat. Chem. Biol.* 9 (2013c) 297–9.
- [81] L. Banci, L. Barbieri, E. Luchinat, E. Secci, Visualization of redox-controlled protein fold in living cells, *Chem. Biol.* 20 (2013b) 747–52.
- [82] L. Barbieri, E. Luchinat, L. Banci, Structural insights of proteins in sub-cellular compartments: In-mitochondria NMR, *Biochim. Biophys. Acta.* 1843 (2014) 2492–6.
- [83] E. Luchinat, L. Barbieri, J. T. Rubino, T. Kozyreva, F. Cantini, L. Banci, In-cell NMR reveals potential precursor of toxic species from SOD1 fALS mutants. *Nat. Commun.* 5 (2014) 5502.
- [84] L. Barbieri, E. Luchinat, L. Banci, Protein interaction patterns in different cellular environments are revealed by in-cell NMR, *Sci. Rep.* 24 (2015) 14456.
- [85] E. Mercatelli, L. Barbieri, E. Luchinat, L. Banci, Direct structural evidence of protein redox regulation obtained by in-cell NMR. *Biochim. Biophys. Acta.* 1863 (2016) 198–204.
- [86] E. Luchinat, E. Secci, F. Cencetti, P. Bruni P, Sequential protein expression and selective labeling for in-cell NMR in human cells, *Biochim. Biophys. Acta.* 1860 (2016) 527–33.
- [87] L. Barbieri, E. Luchinat, L. Banci, Characterization of proteins by in-cell NMR spectroscopy in cultured mammalian cells. *Nat. Protoc.* 11 (2016) 1101–11.

## Highlights

- \* In-cell NMR is very useful to study bio-macromolecules in their natural environment.
- \* Interactions among biological macromolecules in a natural cell can be gained directly.
- \* Artificially high concentration in prokaryotic cells hinders spectral observation.
- \* Best results are obtained with intrinsically unfolded proteins.
- \* The outlook for in-cell NMR of eukaryotic cells is more promising.