In-cell structural biology by NMR: the benefits of the atomic-scale

Francois-Xavier Theillet

In-cell structural biology by NMR: the benefits of the atomic-scale.

Francois-Xavier Theillet*

Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.

ABSTRACT: In-cell structural biology aims at extracting structural information about proteins or nucleic acids in their native, cellular environment. This emerging field holds great promises and is already providing new facts and outlooks of interest at both fundamental and applied levels. NMR spectroscopy has important contributions on this stage: it brings information on a broad variety of nuclei at the atomic scale, which ensures its great versatility and uniqueness. Here, we detail the methods, the fundamental knowledge and the applications in biomedical engineering related to in-cell NMR. We finally propose a brief overview of the main other techniques in the field (EPR, smFRET, cryo-ET…) to draw some advisable developments for in-cell NMR. In the era of large-scale screenings and deep learning, both accurate and qualitative experimental evidences are as essential as ever to understand the interior life of cells. In-cell structural biology by NMR spectroscopy can generate such a knowledge, and it does so at the atomic scale. This review is meant to deliver comprehensive but accessible information, with advanced technical details and reflections on the methods, the nature of the results and the future of the field.

CONTENTS

1. INTRODUCTION

Almost twenty years ago, the last special issue of *Chemical Reviews* about Biological NMR dedicated 5 out of 9 articles to structure determination of proteins or nucleic acids.[1] In the meantime, the field of structural biology has experienced at least two revolutions: the advent of cryo-EM[2,3] and deep-learning assisted prediction[4–6] as effective and dominant techniques. More silent revolutions had also profound impacts, like the ~7,000 structures per year that X-ray crystallography released since 2001. NMR spectroscopy was probably less adapted to such an industrialization and its volume production in the field of structure resolution declined.[7] It supported instead another ground swell in structural biology: the growing awareness about the importance of protein regions lacking a stable fold (the so-called Intrinsically Disordered Proteins, IDPs).[8–10] Altogether, these advances set a new scene, which calls for the next important leap in the field: structural biology shall focus on macromolecules in their cellular environment.[11,12]

Let us consider the average cytoplasmic composition (Figure 1): inorganic ions at about 300 mM,[13,14] total metabolites concentration at 200-300 mM,[15,16] ~200 to 300 g/L of proteins and ~20 to 100 g/L of RNAs (eukaryotic vs bacterial cells),[17–22] ~4 membrane proteins per 100 nm^2 representing ~30% of surface occupancy and mass for the average membranes,[23,24] a phosphorylation site detected on nearly every protein in human cells (4 sites per phosphoprotein on average, most of them at high stoichiometry),[25–28] all these values being naturally cell-, phase-cycle- and organelle-dependent.[13,25–27,29–35] These must have an impact on the conformational behavior of macromolecules, on their accessibility, hence on their binding capacities.[13,36–44] While *in vitro* studies using purified material will always be necessary for an accurate knowledge, a more advanced depiction in the cellular milieu is desirable and, fortunately, attainable now. A number of techniques were adapted or designed with the purpose of achieving structural biology studies on whole cells, among which in-cell structural biology by NMR spectroscopy.

NMR spectroscopy has already had unique contributions that shape our mental representations of biological objects: among others, it permits to characterize the structural dynamics and weak interactions of molecules of (nearly) all sizes in solution; it does so from the picosecond to the hour timescales, in a non-destructive fashion, and reports for information raised directly from atom nuclei, hence at the atomic scale.[45–52] These capacities are transposable to studies in complex media like whole cells. NMR spectroscopy was thus meant to produce valuable knowledge in in-cell structural biology.

In-cell structural biology by NMR emerged in the last 20 years,[53–57] building on the long-term practice of NMR spectroscopists, who manipulate and characterize live cells, tissues and animals in their spectrometers for 70 years already.[58–61 62] Nuclei with a non-zero spin magnetic moment are indeed detectable using NMR spectroscopy in almost every material, in a non-destructive fashion and with a very good penetrance (see MRI for example). The cellular context has nevertheless some inconvenient consequences for the manipulation and detection of nuclear spins.

In-cell NMR has thus its own constraints, methods and capacities, which we will detail in the present review. We will start with a Methods section, where i) we give basic information for non-NMR spectroscopists, and ii) we switch progressively to the core technical details that make studies feasible and valuable or not. Then, we will show that in-cell NMR does not only produce fundamental knowledge, but also has interesting potential in drug research. We will also sketch briefly the current landscape of the emerging field of in-cell structural biology. We tried to deliver beneficial perspectives for in-cell structural biology by NMR spectroscopy.

Figure 1: Model of a bacterial cytoplasm matching experimental concentrations. Macromolecules are shown with a cartoon representation in the foreground, and with a surface representation in the background; metabolites and ions are shown with stick and ball representations, respectively. Adapted from ref 63. Copyright 2016, Yu et al. under the terms of a Creative Commons CC BY license https://creativecommons.org/licenses/by/4.0/.

2. METHODS

To be comprehensive and accessible to the broadest audience, we have to provide some explanations about the theoretical and practical backgrounds of in-cell structural biology by NMR. In-cell NMR inherits of the characteristics of the more standard NMR spectroscopy. Still, in-cell NMR spectroscopy has specific constraints and limits.

We decided to limit ourselves to solution-NMR approaches. Solid-state NMR (ssNMR) provided interesting structural descriptions using native membranes, e.g. on the interfering modes of antibiotics in bacterial membranes,[64–66] or on the structure of prokaryotic and eukaryotic membrane proteins.[67–69] In-cell ssNMR is somehow in its infancy and will certainly be fueled by ^{1}H-detection and Dynamic-Nuclear Polarization (DNP) in the coming years.[70–73] Describing the methods associated to in-cell ssNMR would thus be a more prospective exercise. Importantly, ssNMR is limited to the characterization of frozen cells: it ensures cellular integrity during ssNMR measurements, which require fast Magic-Angle Spinning (MAS, about 10 to 100 kHz) to obtain atomic-scale information on macromolecules. Solution- and ss-NMR report therefore results of different natures.

2.1. Simplified theoretical basis of in-cell NMR.

2.1.1. What type of information.

We must start with some quick reminders for the non-NMR specialists. The basic principles of NMR spectroscopy hold true when applied to cellular samples. In very simplified, rough terms, in-cell NMR spectroscopy can provide information on a population of nuclei i) *via the chemical shifts*: on the chemical context, i.e. the covalent bonds and non-covalent interactions, either intra- or inter-molecular, and more generally on the surrounding chemical structures; ii) *via the chemical shifts and the dipolar interactions*: on the local conformations and the intra- or inter-molecular distances between nuclei; iii) *via the chemical shifts, the dipolar interactions and the magnetization relaxation*: on the local conformational dynamics, on the molecular tumbling and the diffusion coefficients, and on interaction kinetics. The readers can find more background information in the attached references.[52,74–79]

The common NMR techniques can report for structural and interaction dynamics ranging from the nanosecond to the second time scales; Recording time series of NMR spectra permits to monitor events occurring in the minute to hour time scales. Hence, it is probably difficult to make a comprehensive list of phenomena that in-cell NMR can help to investigate. We can cite conformational dynamics, folding/unfolding equilibria, intermolecular interactions, chemical modifications, intracellular mobility, membrane permeability (Figure 2).

Small and large molecules can both be investigated, even though their intrinsic tumbling times provoke very different relaxation rates, hence different signal intensities per molecule (the larger the molecule, the lower the signal).

2.1.2. The isotope filter.

NMR spectroscopy permits the selective observation of a chosen type of isotope in a sample, which we call the isotope filter for convenience in this article. Let us give an explanation in three sentences for the non-NMR experts: NMR spectroscopy detects the precession of nuclear magnetic moments, which adopts their resonance frequency according to the Lar-

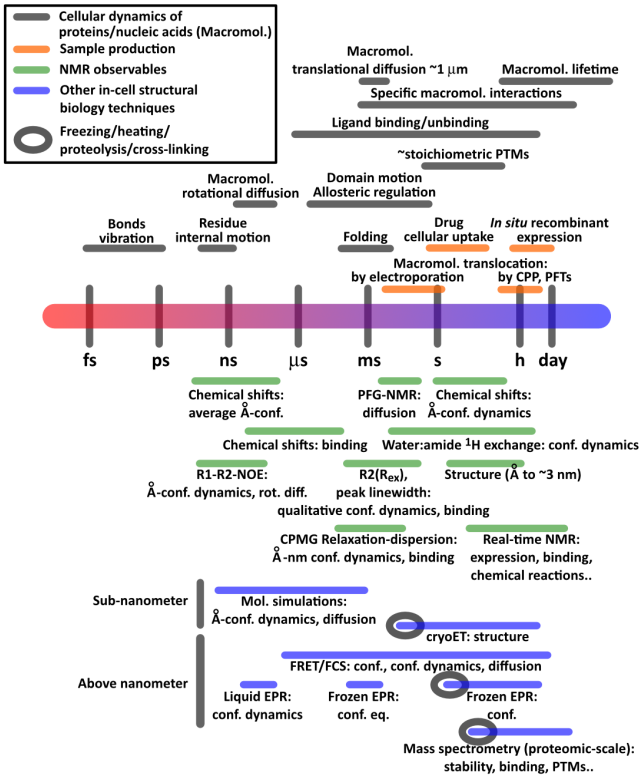


Figure 2: Timescales of i) cellular events and dynamics for proteins/nucleic acids (named "Macromol.") in grey; ii) cell delivery of drugs or macromolecules to generate in-cell NMR samples (CPP: cell-penetrating peptides; PFTs: pore-forming toxins) in orange; iii) structural information that in-cell NMR observables report (conf.: conformation) in green; iv) structural information that other in-cell techniques can provide in blue. Circles indicate the current freezing timescales for cryo-electron tomography (cryoET) and Electron-Paramagnetic Resonance (EPR), or the heating/proteolysis/cross-linking procedures used for later analysis by Mass-Spectrometry. We distinguished atomic/residue-scale structural information (structure) from nanometer scale conformational analysis (conf.). The typical cell dimensions range from 1 to ~40 µM for bacterial and mammalian cells, respectively. Frog and zebrafish oocytes are ~1 mm large.

mor equation $\omega = -\gamma_{nucl.} * B_{loc}$, with $\gamma_{nucl.}$ the gyromagnetic ratio of the observed nucleus, and B_{loc} the local magnetic field. Every isotope has a unique $\gamma_{nucl.}$, which provokes its precession in a defined range of frequencies modulated by the variations of B_{loc} felt by the nucleus, usually in the p.p.m. range. Hence, one can manipulate and observe a certain isotope selectively, by tuning the probe circuitry to its corresponding NMR frequency window -the rationale is actually the same than picking a radio station by its frequency.

However, to be NMR-visible, an isotope has to have a non-zero spin magnetic moment. At first glance, this could ruin the enthusiasm about in-cell NMR: the most abundant isotopes of carbon (^{12}C, natural abundance 98.9%), oxygen (^{16}O, 99.8%) or sulfur (^{32}S, 95%) have a null spin. Moreover, ^{14}N (99.6% of nitrogen) has unfavorable spin and quadrupolar moment values, which makes it poorly detectable with the current NMR techniques (N.B.: Except for the small, symmetric molecules NH_4^+ and NO_3^- or using ssNMR approaches still far from routine and from relevant cellular concentration ranges)[80,81]. In-cell NMR spectroscopists turned this drawback into a blessing: by delivering in cells molecules enriched in ^{13}C and ^{15}N,

two isotopes with NMR friendly characteristics, they can execute ^{13}C- or ^{15}N-editing NMR techniques in an almost blank cellular background (Figure 3). In this regard, non-natural ^{19}F-containing amino acids or nucleic acids can also be convenient, because cells do not contain any fluorine.

Finally, it is still possible to detect selectively a molecule of interest without any fancy isotopic labeling: protons can be directly observed by ^{1}H-NMR and specific molecules can be nailed and characterized if their signals surpass those of the cellular background. This works either for i) small ligands, whose sharp, intense signals are straightforward to recognize; ii) proteins containing chemical functions showing very peculiar chemical shifts out of the standard amide, aromatic and alkyl regions.

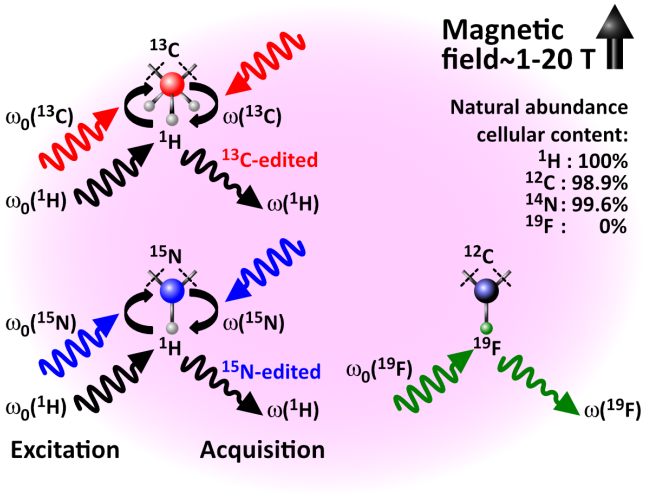


Figure 3: Operating principle of the NMR isotope filter. Under strong magnetic fields, electromagnetic pulses (sinusoidal arrows) excite populations of isotopes selectively according to their respective NMR frequency ω_0. ^{1}H and ^{19}F are excited and detected in preference, because their high gyromagnetic ratio ensures the best sensitivity. Subsequent ^{1}H/^{13}C/^{15}N pulses provoke i) the magnetization transfers through covalent bonds, ii) the sampling of the residue-specific $\omega(^{13}C/^{15}N)$, and iii) the back-transfer to ^{1}H nuclei for the final ^{13}C/^{15}N-edited ^{1}H-NMR acquisition (decaying sinusoidal). Broadband ω_0-pulses permits the excitation of large populations of the chosen isotope nuclei. The acquired signals is a sum of resonances at residue-specific ω-frequencies, which are deconvoluted via Fourier transformation. An in-cell NMR sample is generated from ^{13}C/^{15}N/^{19}F-labeled macromolecules either i) delivered in cells grown in natural abundance culture media, or ii) overexpressed in presence of ^{13}C/^{15}N/^{19}F-labeled amino acids/nucleotide precursors.

2.1.3. Which nuclei and molecules do we observe? We have to introduce a number of principles to explain the (un)capacities of in-cell NMR.

First, the higher the gyromagnetic ratio, the better is the potential sensitivity: if we neglect the relaxation effects, $S/N \propto \gamma_{excitation} * \gamma_{detection}^{3/2}$. The nuclei of interest for in-cell structural biology line up in this order: $\gamma(^{1}H) \sim \gamma(^{19}F) \sim 2.5*\gamma(^{31}P) \sim 4*\gamma(^{13}C) \sim 10*\gamma(^{15}N)$. This is the reason why ^{13}C- or ^{15}N-editing NMR techniques start preferably by exciting protons and then transfer the magnetization to vicinal ^{13}C or ^{15}N nuclei. In solution NMR, transferring back the magnetization to protons yields the best sensitivity in most cases. In ssNMR, protons build up strong dipolar interactions, which are difficult to cancel; protons' ssNMR is thus unpractical, unless very fast Magic-Angle Spinning (MAS) (>60 kHz, >100 kHz for fully-protonated samples) is used.[82] This requires pieces of equipment that are not yet widespread, and in-cell ssNMR studies is thus carried out using ^{13}C-detection so far.

Second, the slower the tumbling time, the faster is the loss of magnetization coherence (measured as the T2 relaxation) and thus the weaker is the observed signal (Figure 4). This phenomenon affects to higher extents the high gyromagnetic ratio nuclei. The cellular milieu tends to slow down molecular tumbling, because of multiple specific and unspecific interactions. Hence, in-cell solution NMR has allowed the observation of globular proteins below 30 kDa, and the in-cell NMR methods used currently will hardly permit the observation of species above 50 kDa. Advanced strategies exist for *in vitro* investigations of large complexes up to MDa sizes though, which are made possible notably by protein deuteration.[83–85] The NMR effective size is that of the molecular complex, except for IDPs, whose non-binding regions remain highly mobile. In contrast, ssNMR is theoretically less limited by size, because MAS cancels out the effects of slow tumbling to a large extent. However, peak linewidths are traditionally large in ssNMR, and larger molecular sizes come with more nuclei: this yields crowded spectra with overlapping signals. Thus, in-cell ssNMR requires to achieve selective isotope labeling schemes (e.g. amino acid specific labeling), and/or multiple magnetization transfers between ^{13}C- and ^{15}N-labeled amino acids to edit very well-defined spin systems.

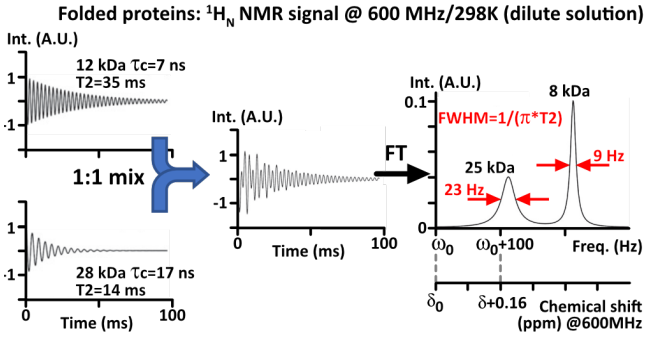


Figure 4: T2 relaxation of folded proteins depends on molecular tumbling, or on the related correlation time τc (the time necessary to rotate through an angle of one radian), which is commonly quantified via the T1 and T2 NMR relaxation times. In the present model situation, two proteins are mixed at a 1:1 ratio, which yields an NMR signal (FID: Free Induction Decay) equal to the sum of both signals. This is deconvoluted via Fourier Transform (FT) to extract the resonance frequencies. A slow-tumbling molecule has a fast T2 relaxation, which yields broad, low intensity peaks upon FT. The FT process generates indeed Full Width at Half Maximum (FWHM) inversely proportional to the T2 relaxation, but the integral of the peak is proportional to the quantity of molecule. This simple situation corresponds to a single pulse experiment, followed by the FID acquisition. Two-dimensional ^{1}H-^{15}N or ^{1}H-^{13}C NMR experiments use multiple pulses and delays during the pulse sequences, which last ~20-100 ms: only a few percent of the starting coherent magnetization is left for the FID acquisition. We present here the ideal NMR signals of folded proteins in absence of any sample inhomogeneity, any binding or conformational exchange, or any water/amide proton exchange. Importantly, protein deuteration slows down T2 relaxation. Adapted from ref [86]. Copyright 2011 Elsevier.

Third, interactions with on/off rates in the µs/ms time scale can provoke important loss of magnetization coherence and thus drastic signal losses, even if only a low proportion (5-10%) of the molecule is in interaction at a given time point (Figure 5). This effect depends on a number of parameters (the on/off rates, the free and bound populations, the chemical shift changes upon binding, the size of the complex). Conformational exchange can have the same effects if it occurs at the same time scale. These phenomena are subjects of very interesting NMR investigations *in vitro*, but these are operated at millimolar concentrations of the observed molecules, which is not compatible with in-cell studies.[51,87–89] In-cell ssNMR is not concerned by these considerations, because it is usually performed at freezing temperature.

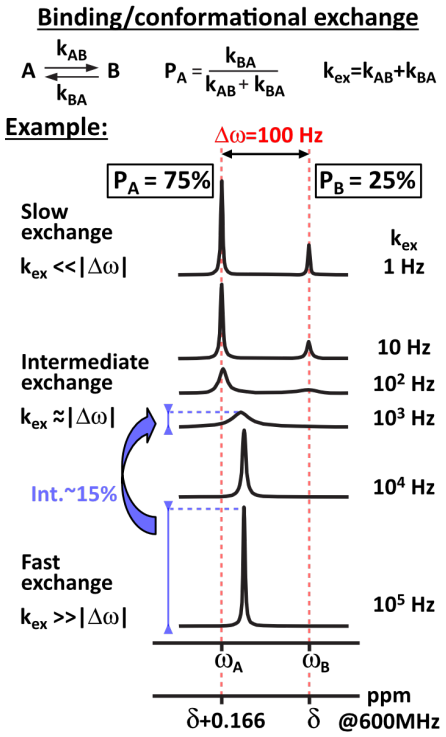


Figure 5: Effects of binding exchange or conformational exchange on NMR signals. Here, we present a simple two-state equilibrium, where A and B are i) either the free and bound states of a molecule, ii) or two conformations of the same molecule; we fixed the chemical shift difference between these two states at 100 Hz and the population P_A:P_B=0.75:0.25. If the exchange between A and B is slow (1 s timescale), two peaks are observed: the two populations generate two resonances that are deconvoluted by Fourier Transform (see Figure 4). If the exchange is fast (sub-µs timescale), the resonances of the two states are averaged during the acquisition. If exchange occurs in the chemical shift timescale (µs-ms), it provokes line-broadening and lower peak intensities; it also manifests itself as a supplementary apparent T2 relaxation. This line-broadening is modulated by the chemical shift difference between the two states, the populations of the two states, the exchange rate, and the intrinsic T2 relaxation of the two states. Techniques exist to quantify all these parameters from T2 measurements at different spin-echo frequencies (see chapters 2.3.10. and 3.1.2.). Adapted from ref [79]. Copyright 2009 Elsevier.

Fourth, NMR spectroscopy reports signals from a population of equivalent nuclei. The peak linewidths depend on the relaxation that we evoked above, but also more simply on the homogeneity of the observed population. Hence, only very broad peaks would emerge from equivalent nuclei, whose population would experience too many diverse chemical environments, e.g. steady interactions with too many partners, or too many conformations. The signal intensity of a population of equivalent nuclei is proportional to the inverse of the peak linewidth: broadening due to multiple interactions of conformations can ultimately result in peak disappearance. This is more likely to occur for ssNMR because of the many states that can be trapped in a frozen state.

Fifth, NMR spectroscopy reports the addition of signals from all the nuclei inserted in the probe, and cannot inform about their localization at the submicrometer scale, i.e. the subcellular scale. We have seen above that signals from certain populations of nuclei can be lost, often because of interactions. If a subpopulation of the molecules of interest sticks to a membrane for example, it does not contribute to the observed signal in solution NMR. More generally, if populations of the molecules of interest behave differently in the many subcellular compartments, NMR spectroscopy will provide averaged information.

Altogether, the integral of the observed NMR signals are proportional to the corresponding populations of nuclei. This makes NMR spectroscopy a quantitative technique. However, a number of cellular events can affect the emerging NMR signal intensities. Hence, the risk is high that in-cell NMR spectra report only for a "NMR-visible" subset of the whole intracellular population of the studied molecule. Hence, proper conclusions require verifications from orthogonal techniques i) to quantify the total number of observed molecules (e.g. western-blots, quantitative mass-spectrometry) and ii) to localize the observed molecules (e.g. fluorescent microscopy). It is also advisable to record NMR spectra from broken cells, where inhomogeneities are removed and molecules are diluted: this permits to recover relaxation behaviors close to the *in vitro* ones obtained using purified, dilute material. Ideally, the quantification of the detected nuclei would also require the measurement of their NMR relaxation rates, in order to evaluate the lost magnetization through the pulse sequences.

We enumerated a number of constraints and possible bias of in-cell NMR observation. These can hamper certainly the investigation of a number of molecules. We want to stress also the positive aspects: these restraints leave an immense field of possible investigations, as we shall see in the rest of the manuscript.

2.2. Sample production

2.2.1. Strategies and isotope labeling

We discuss in the following the first steps of in-cell structural biology investigations executed by NMR spectroscopy, i.e. the production of samples designed for the observation of either ligands or of macromolecules. Cells are eventually entrapped in gel matrices and supplied with a constant flow of fresh medium (see chapter 2.2.11.). The ligand-oriented NMR studies use cells grown in standard conditions. At the opposite, macromolecule-oriented studies rely often on the observation of isotope-labeled material in a non-labeled or poorly labeled cellular background, which is much less standard. This can be achieved in two ways (Figure 6).

The first possibility it to perform the direct production of the molecule of interest in the studied cells, which we call "*in situ* production" for convenience. The second strategy follows a two-step rationale: the initial production is obtained from

recombinant expression or synthesis, and the purified material is later injected into the chosen cells grown on natural abundance media. We can call it the "delivery approach". A better NMR contrast is to be expected from this second approach: natural abundance cellular content yields little background noise once the NMR isotope-filter is executed. At the opposite, *in situ* recombinant production of the protein(s) of interest implies the simultaneous incorporation of the chosen isotope-labeling into the whole cellular content; this calls for robust expression of the studied protein to surpass endogenous NMR signals, or to limit the production of endogenous proteins. This approach has the enormous advantage to avoid purification steps and the unfortunate, but common troubles met with purified proteins' stability and solubility.

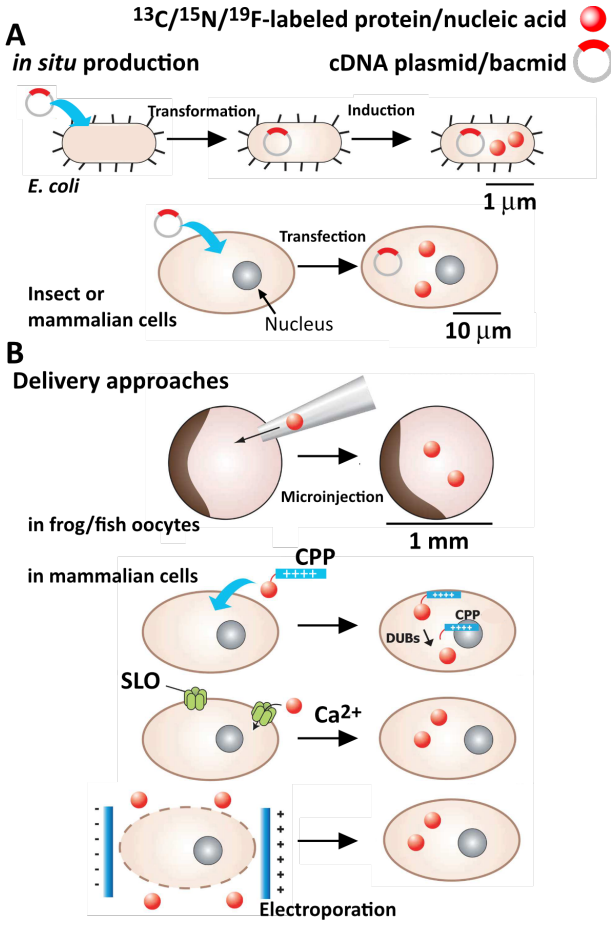


Figure 6: Cartoon representation of the two types of in-cell NMR samples. **A)** In the "*in situ* production" approach, cells are transfected with plasmids encoding for the protein(s) of interest, and the expression of the protein of interest is achieved in a medium supplemented with $^{13}C/^{15}N/^{19}F$-labeled amino acids or nucleotides, or precursors thereof; in the case of *E. coli* cells, protein production is induced at a chosen timepoint; in the case of insect cells, the bacmid transfection leads to baculovirus assisted protein production; in the case of mammalian cells, transient transfection was used in most cases, but permanent transfection is also possible (eventually with inducible expression); **B)** In the "delivery approaches", $^{13}C/^{15}N/^{19}F$-labeled, purified proteins are transduced in cells, either using microinjection for oocytes, either using cell-penetrating peptides (CPPs), the pore-forming toxin Streptolysin O (SLO), or electroporation; CPPs can included in chimera constructs together with ubiquitin and separated by deubiquitinases (DUBs); SLO pores are "resealed" by short exposure to Ca^{2+}. Adapted from ref [90]. Copyright 2020 Elsevier.

In both cases, the production of the protein or nucleic acid of interest is carried out with the supplementary purpose of isotope-labeling. Presenting all the refined aspects of this topic is beyond the scope of our review. Still, we can mention some recent references focusing on the recombinant production of proteins dedicated to NMR analysis: i) for a uniform ^{2}H-,^{13}C and/or ^{15}N-labeling in *E. coli*,[91–95] yeast,[96,97] insect cells,[98–101] or mammalian cells,[102,103] ii) for ^{13}C-methyl-labeling in *E. coli*,[104,85,105] yeast[106–108] or insect cells,[109–111] iii) for amino acid specific labeling in *E. coli*,[112–114] insect cells,[115] iv) for unnatural amino acids labeling containing ^{19}F-moieites [116–119] or other non-natural chemical functions.[119,120] The so-called cell-free approaches also offer a range of isotope labeling possibilities.[121–125] Concerning the production of nucleic acids for NMR, methods exist and are continuously being developed, using enzymatic reactions for uniformly or nucleic acid-specific ^{2}H-,^{13}C and/or ^{15}N-labeled nucleic acids,[126–131] or solid-phase synthesis allowing position-specific labeling,[126,128,131–133] or the insertion of non-natural bases like those incorporating ^{19}F-moieties.[134–139] The ^{13}C-methyl-labeling of DNA has also been proposed recently.[140] The field is vast though, and we guess that many of the non-cited labeling schemes would find interesting applications for in-cell NMR studies. We elude the questions related to NMR signal assignment.

2.2.2. *In situ* recombinant production in *E. coli.*

The first in-cell structural biology studies by NMR were driven using recombinant, overexpression of labeled proteins and subsequent observation of the whole cells. This corresponds to the core expertise of most of the structural biology laboratories, and it was the natural entry to in-cell studies.

To start with, we must mention a pioneering report in the mid-1990's from Gronenborn and Clore, who noticed that ^{15}N-labeled proteins overexpressed in *E. coli* (grown in a minimal medium supplemented with ^{15}N-ammonium) could generate decent ^{15}N-edited spectra in crude cell extracts. Next, Brindle and colleagues incorporated 5-fluorotryptophan in auxotrophic strains of *Saccharomyces cerevisiae* transformed to overexpress enzymes, among which two were detected by ^{19}F-NMR of the whole cells, while two others became detectable only after cell lysis.[141,142]

From 2001, Dötsch and colleagues were the first to exploit the recombinant overexpression in *E. coli* in the context of a structural analysis in cells.[143] We assume that recombinant expression in *E. coli* is well-known from the reader: it is the expression organism of >75% of the protein structures in the PDB. Dedicated reviews exist on the subject.[144,145] Among all their practical aspects, *E. coli* strains dedicated to protein expression allow cost-effective isotope-labeling (see the cited protocols above). The most popular strains (e.g. BL21-(DE3) strains) express the phage T7 polymerase. It permits high and inducible expression of plasmid coding regions under control of the T7 promoter, yielding up to 50% of the final cellular protein content.[144,145] Dötsch and colleagues recorded exploitable ^{15}N-edited solution-NMR spectra using a *E. coli* cell slurry obtained from 50 mL of isotope-labeled culture that expressed a protein of interest.[146] They estimated that this protein had to represent more than ~5% of the total soluble proteins to provide useful NMR signal, well-above the cellular background.[147] This corresponds to 0.15 to 1 mM intracellular concentrations.[148] This is fortunately the order of magnitude reached in *E. coli* overexpression systems, i.e. 0.1-10 mM of

intracellular recombinant protein, depending on the plasmid, the encoded protein, the temperature during expression and the *E. coli* strain.[149,150] However, such an in-cell solution-NMR visibility rule holds true only in cases where the protein tumbling is poorly affected by the cellular environment (see Chapters 2.1.3. and 2.3.7.). Dötsch and colleagues explored the possible schemes and timing of cell growth and expression to obtain the highest NMR signal of interest and the lowest signal from the cellular background. Expression was induced in a medium supplemented with $^{15}NH_4^+$ as sole source of nitrogen, while the cells were in exponential growth phase. It appeared that growing cells in natural abundance Luria-Bertani (LB) medium before expression in the ^{15}N-medium did not change drastically the background signal, neither did the supplementation with rifampicin that inhibits the transcription of endogenous proteins.[147] In contrast, Baldus and colleagues have shown recently how rifampicin supplementation is important to clear the spectra recorded using in-cell ssNMR spectra,[72,151] which detects all the large and abundant, endogenous macromolecules that in-cell solution-NMR miss (e.g. chaperones, ribosomes, membrane proteins, peptidoglycans, …). Dötsch and colleagues removed also convincingly the cellular signal background by using ^{13}C-depleted glucose (99.9% ^{12}C) as sole source of carbon during *E. coli* growth before inducing protein expression;[152] they also advocated and exemplified the benefits of combining amino acid specific labeling schemes that permit ^{13}C/^{15}N editing, e.g. [1-^{13}C]-isoleucine and ^{15}N-labeled amino acids.

In-cell solution NMR protocols from Dötsch and colleagues were reused by the solution-NMR groups of Pielak,[153–164], Hubbard,[165] Alexandrescu,[166] Ito,[167–169] Crowley,[170] Banci,[171,172] Gierasch,[173] Christodoulou and Waudby,[174,175] Selenko,[176,177] Felli and Pierattelli,[178] Cowburn,[179] Pastore,[180] and Yao.[181] Shekhtman and colleagues proposed to upgrade this approach with the purpose of allowing the characterization of protein:protein interactions.[149,182] They established protocols to overexpress sequentially two proteins encoded in plasmids with orthogonal antibiotic resistances and induction triggers: the first protein expression was induced in a ^{15}N-labeled medium under the control of a first gene repressor; the second was expressed in a ^{14}N-containing natural abundance medium under the control of the T7 polymerase system, while repressing the first gene.[149] Hence, only the first protein was ^{15}N-labeled and detectable in ^{15}N-edited spectra, which simplified the analysis and avoided mislead interpretation of NMR peaks shifting.[149,182] Shekhtman and colleagues reused and eventually complexified their approach with a third plasmid coding for a third protein.[183–187]

In all these "in-*E. coli*" solution-NMR studies cited above, ^{15}N-labeling has been preferred because it generates much less cellular background in ^{15}N-edited spectra than ^{13}C-labeling in ^{13}C-edited spectra.[146,147] However, specific labeling by ^{1}H/^{13}C-isoleucines,[174] or ^{1}H/^{13}C-methionines -and to a lesser extent, ^{1}H/^{13}C-alanines-[148] was shown to produce isolated signals in relatively empty regions of 2D ^{1}H-^{13}C NMR spectra, provided that the examined cells had grown in D_2O.[148] They also showed the feasibility and usefulness of ^{15}N-lysine or ^{15}N-histidine labeling for in-cell studies.[146,147,188] Combined ^{15}N-^{13}C-labeling were carried out for in-cell NMR assignment and structural analysis,[179,167–169,178,175] also accompanied with deuteration and methyl labeling for in-cell structure determination by Ito and colleagues.[167–169] Not surprisingly, deuteration diminished the linewidth of in-cell NMR signals (Figure 7).[147,149,189,190]

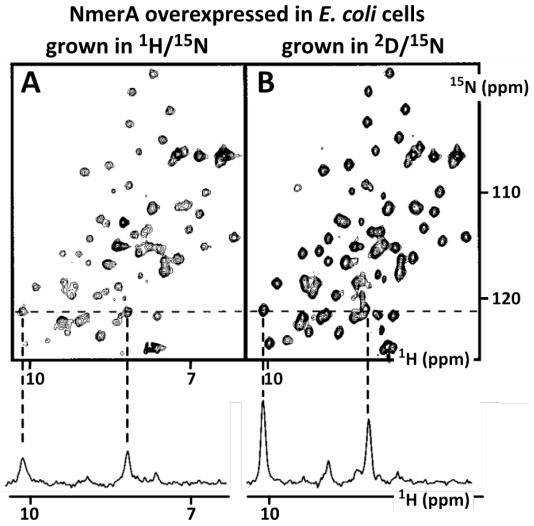


Figure 7: Effects of deuteration on in-cell samples produced via recombinant overexpression in *E. coli*. **A)** 2D ^{1}H-^{15}N spectrum of the model protein NmerA expressed in ^{15}N-labeled and **B)** in ^{2}D- and ^{15}N-labeled culture media. Adapted from ref [147]. Copyright 2001 American Chemical Society.

The incorporation of ^{19}F-containing unnatural amino acids has also been carried out with success for in-cell solution NMR studies. Mehl and colleagues proposed initially to incorporate trifluoromethyl-phenylalanine via amber recoding in a position-specific fashion.[191] Since 2010, Li, Pielak and their colleagues have used simpler approaches by expressing proteins in a medium supplemented with 3-fluoro-tyrosine, [192–195], and 6-fluoro-tryptophan[195,196]or 5-fluoro-indole,[197–200] which are integrated by *E. coli* cells in proteins as tyrosine or tryptophan, respectively. Crowley and colleagues have indeed shown that 5-fluoro-indole are readily utilized by *E. coli* cells as a precursor of tryptophan.[116,201]

Importantly, Pielak and colleagues have recently used the Tuner™ *E. coli* strain, because it permits a homogeneous, IPTG-concentration dependent expression in the cell population.[150,202] The lactose permease has been deleted in Tuner™ cells, which makes the whole population equally permeable to IPTG. At the opposite, IPTG penetrates the more classical BL21(DE3) in a cell-specific fashion. Hence, at [IPTG]<0.25 mM, the induction is stochastic and binary, yielding cell populations, among which individuals either overexpress the recombinant protein of interest, either not at all.[202] Future studies focusing on protein intracellular interactions should exclusively use *E. coli* strains like Tuner™, whose recombinant expression is tunable, but more importantly homogeneous.

To finish with *E. coli* samples, we must give a few important comments. First of all, when settled as a cell slurry in a NMR tube, they have a limited lifetime of about 4-6 hours:[203] Lysis and leakage occur soon and a number of early contributions were biased by the measurement of proteins leaked in the supernatant.[204,205] It is also important not to store and freeze *E.coli* cells before solution NMR measurements to avoid cell lysis.[193] In-cell solution NMR analysis of *E. coli* samples has been popular because of its straightforward accessibility, but its attractivity declined because of the common low-quality of the obtained spectra: *E. coli* cells provoke important signal broadening, often beyond detection. This is notably due to

high intracellular density associated to high viscosity and countless unspecific interactions. In this regard, some practical observations should be reminded: i) thicker cell slurries generate broader linewidth, notably because of lock and shim troubles linked to ^{2}H-water detection; [146,147,206] ii) protein induction at different cell stages yields spectra of astonishing different qualities.[206] In our opinion, this probably means that a number of key aspects might have been overlooked. NMR spectroscopists tend to apply systematic protocols for overexpression using the T7 system, but *E. coli* cells reacts in many different ways to heterologous protein production.[207] The T7 polymerase system and its induction by IPTG are known to provoke stress and metabolic burden,[208–210] which materialize at the proteomic,[211] transcriptomic,[210,211] and even genomic levels.[212] *E.coli* cells content does of course vary at different growth phases, but it is probably as important to realize that the cell population is neither homogeneous nor constant upon IPTG induction: i) expression upon IPTG induction is highly heterogeneous in a population of (DE3) *E. coli* cells;[150,202,212] ii) some cells survive to the heterologous expression, whereas some other cells do not, which gives rise to important natural selection effects upon incubation with IPTG.[212] It would thus be advisable to use strains that permit tunable control of protein expression in the future (more formally: protein expression should be a continuous function of the inducer concentration in every cell). It will also be important to pay attention to the growth phase, the metabolic state and the population homogeneity of the cells, both for spectroscopic purposes and for the biological interpretation.

2.2.3. *In situ* recombinant production in yeast.

Heterologous expression in *E. coli* is of course limited in terms of biological relevance for the analysis of eukaryotic proteins (e.g. native chaperones, transcription machinery and co-translational processing, cellular organization, metabolites, etc…). Eukaryotic systems have also been tested for in-cell structural biology by NMR spectroscopy.

We mentioned earlier (chapter 2.2.1.) already the first reports from Brindle and colleagues using heterologous expression and ^{19}F-tryptophan incorporation in *S. cerevisiae*.[141,142] Recombinant expression in *E. coli* took over in the years 2000's (Chapter 2.2.1.), before yeasts were reused for in-cell structural NMR purposes a few years ago. These present interesting features: i) their ease of genetic manipulation, ii) their importance in biopharmaceutical protein production,[213–215] iii) their capacity to integrate the cost-effective ammonium and various carbohydrates as sole source of nitrogen and carbon for isotope labeling, iv) their numerous amino acid auxotroph strains, v) their established impact in the field of protein:protein interactions discovery, with the so-called yeast two hybrid technique.[216] In-cell NMR could provide the atomic-scale description counterpart of the protein:protein binding revealed by yeast two hybrid. Shekhtman and colleagues tested *Pichia pastoris* to carry out in-cell NMR studies, and have shown that heterologous expression must be carried out using a medium supplemented with methanol and not with dextrose: the second situation provokes the sequestration of the proteins of interest in vesicles, where they become undetectable by solution NMR.[217] The authors claimed that the association with cellular RNA was responsible of NMR disappearance.[217] Wall and Hough managed to record well-behaved spectra of a disordered protein at an intracellular concentration of ~175 µM in *S. cerevisiae*, expressed in a medium supplemented with $^{15}NH_4^+$ and ^{13}C-galactose. Interestingly, these cells were very stable and no leakage was observed in more than 50 hours. To the best of our knowledge, these are the only three reports of in-cell NMR in yeast. This is probably linked to the fact that only few NMR laboratories use yeast for standard recombinant protein production, which requires equipment and a certain expertise. Moreover, quiescent yeast tend to reorganize their cytosol and to produce phase-separated storage organelles.[218,219] The heterologous protein of interest might shuttle in an out of these compartments, which would result in an important heterogeneity of cellular environments: this is not compatible with a proper NMR analysis, which reports averaged signals (see Chapter 2.1.3.).

2.2.4. *In situ* recombinant production in insect cells.

Like yeast systems, insect cells have not been extensively used for in-cell NMR studies. The necessary set up and expertise are not widespread in the NMR community, except for laboratories working on GPCRs dynamics (see for example [220–223,111]). Moreover, the metabolic capacities of insect cells are limited and this implies the use of culture media supplemented with amino acids, whose ^{2}H/^{13}C/^{15}N versions are costly. Hence, isotope labeling of recombinant proteins in insect cells has been traditionally expensive, even though cost-effective solutions have been proposed lately.[98–101 109–111 115] Insect cells are nevertheless the second production organism quantitatively in the PDB, notably for multiple protein complexes and human therapeutic targets (e.g. ~85% and 50% of GPCRs and Ser/Thr/Tyr kinases structures in the PDB, respectively)[224,225]. Ito and colleagues have expressed five proteins in insect cells using the baculovirus approach, and showed that they obtained samples amenable for in-cell solution NMR analysis.[226,227] In this system, the cytosolic proteins of interest reach intracellular concentrations of ~100-200 µM, and 2.10^7 cells are enough for one in-cell sample, which require only 5-10 mL of culture medium. This made it accessible to characterize proteins in cells using various isotope labeling schemes, either uniform or amino acid specific (A/I/L/V/T/F/W/Y),[227] although these would have frightening prices per liter, between 1000 and 4000 euros per liter.[98–101,110,115] These can be achieved using depleted culture media supplemented with isotope-labeled amino acids. Similar to what has been observed in *E. coli* cells, ^{15}N-labeling generates less cellular background signal in ^{15}N-edited spectra than ^{13}C-labeling in ^{13}C-edited spectra. [226,227] Insect cells remain >90% viable 8 and 24 hours without and with constant medium replenishment, respectively.[227] The current cell lines have indeed been engineered to stand the baculovirus infection and not triggering apoptosis.[228,229] However, the infection has profound consequences on insect cells, affecting their metabolism, shutting off their endogenous protein synthesis (except that of chaperones), rearranging their cytoskeleton, and stopping the cell cycle in G2/M phase.[230,231] The viral infection is moreover intrinsically not extremely reproducible. Transient transection systems have been proposed, which may preserve more "normal" cellular states that would better fit to the ambition of characterizing proteins in the native environment.[232]

2.2.5. *In situ* recombinant production in mammalian cells.

A majority of structural biology studies by NMR focuses on mammalian proteins.[233] Hence, NMR investigators would naturally prefer to execute their studies in mammalian cells. Such studies exist, but most of them have been carried out

from samples generated upon cell delivery of isotope-labeled proteins (see chapter 2.2.3.).

To be exhaustive, we shall mention the very first in-cell studies of proteins, published in 1975, which actually did not require recombinant, overexpression: erythrocytes and a contain hemoglobin at ~5 mM (~350 g/L), which was labeled by feeding mice with L-[2-^{13}C]-histidine. Because hemoglobin is by much more abundant than any other histidine-containing cellular molecule, NMR analysis of erythrocyte suspension provided direct information on hemoglobin and its intracellular tumbling time.[234] Similarly, ^{15}N-edited spectra of hemoglobin in cells were obtained in 1976 using cancer cells from erythrocyte precursors (namely Friend leukemia cells)[235] grown on a medium supplemented with ^{15}N-glycine.[236]

The Banci group undertook the challenge to express their proteins of interest *in situ*. With the help of cell biology experts, they established a set up to produce isotope-labeled proteins using transient transfection of the widespread HEK293 cells.[237,238] The high expression vector pHLsec from the Aricescu lab[239] was instrumental in their success: this plasmid includes a so-called "CAG promoter" ensuring high levels of protein production; its secretion sequence was removed during the cloning steps.[238] HEK293 cells are commonly used in laboratories because of their ease of manipulation and transfection, their high capacities in protein production associated with human glycosylation patterns.[240–242] Transfected HEK293 cells do not permit the industrial months-long cultures and the yields obtained with monoclonal Chinese Hamster Ovary (CHO) cell lines, up to several grams per liter of fed-batch culture for secreted antibodies.[243–246] However, HEK293 cells produce commonly about 1-10 mg of recombinant protein per liter of culture in the 48-96 hours following their transfection,[240,239,247–250] which was instrumental in the rapid description of the SARS-CoV2 spike protein and its recognition by antibodies for example.[251–254] This is in the order of magnitude reached by Banci and colleagues, who reported intracellular concentrations between 50 and 300 µM for various proteins,[172,237,238,255–261] 48 hours after transfection: this corresponds to a few mg per liter of culture, knowing that they incubate ~10-30 million adherent cells (1-2 pL per cell) in 20 mL of culture medium.[237,238] Because wet pellets of cells are poorly compact, this number of cells is enough to fill a 3 mm tube (~130 µL). Hence, although the commercial (^{13}C/)^{15}N-isotope labeled media cost up to 10,000 euros per liter, in-cell NMR samples made of uniformly ^{15}N-labeled, or ^{15}N-labeled cysteine HEK293 cells are still affordable (NB: Banci et al. used polyethylenimine, PEI, a costless transfecting agent). A home-made algal autolysate medium was proposed for uniform ^{15}N-labeling, which reduced the costs by a factor 3, but also the protein yields by a factor 2.[103] Because cultured mammalian cells are not capable of synthesizing a number of essential amino acid must, other amino acid specific isotope labeling could be tested with profit.[262–265] The protein concentration can be tuned by varying the transfected plasmid carrying the coding DNA, and the combined transient transfection of two plasmids has been used successfully to express two proteins concomitantly.[237,238] This is notably useful when the protein of interest requires the expression of a cognate partner or chaperone at equivalent levels.[172,237,238,255,257] This results however in the expression of two isotope labeled proteins, and thus in the addition of NMR signals from the two proteins in congested spectra. To avoid it, Banci and colleagues have generated stable transfected cells expressing a first protein in a natural abundance medium, which can be silenced by shRNA treatment while transfecting a plasmid encoding a second protein expressed in an isotope-labeled medium.[266] Finally, when settled as a pellet in a NMR tube, the common cultured mammalian cells exhaust their ATP in only 1-2 hours and remain viable >90% for 2-3 hours, but they do not leak for 12-20 hours at low temperature (<283K); alternatively, these cells can be kept viable for days under constant medium replenishment using a flow-probe bioreactor, which is closer to native cellular conditions.[57,267–269]

We shall finish this section with three remarks. First, proteins expressed *in situ* can also be observed without isotope labeling if they can be detected by intense NMR signals in singular, background-free spectral windows. This has been exploited by Banci and colleagues: ^{1}H-NMR signals from histidine side chains of the enzymatic center of carbonic anhydrase 2 (CA2) resonate between 11 and 16 ppm in a very favorable fashion.[260,261,269] A high intracellular concentration is here sufficient for being detected by in-cell NMR (they used ~150 µM, but ~20 µM would be enough). Katahira and Trantirek groups have used a similar approach to detect the in-cell ^{1}H-NMR signals of nucleic acids (10-20 µM intracellular concentrations using cell delivery, see below).

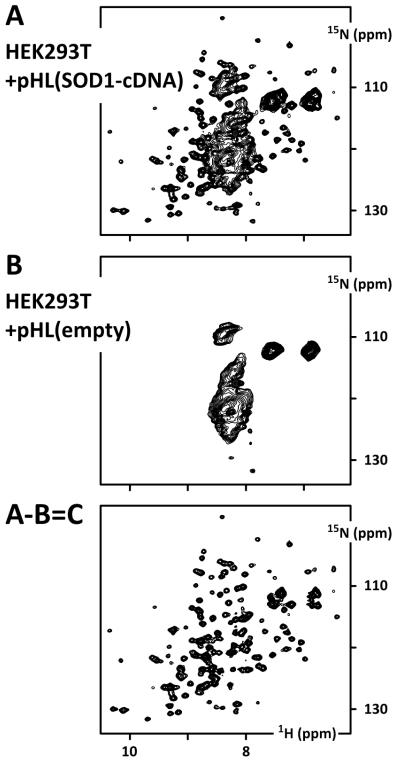


Figure 8: Subtraction of the cellular background NMR signal. HEK293T cells were cultured in a medium containing ^{15}N amino acids after transfection by an expression plasmid pHL containing **A)** a DNA insert encoding SOD1 or **B)** no DNA insert. The resulting samples generate two in-cell spectra, which can be subtracted to obtain a "background-free" spectrum (**C**). Both samples have to be cultured in the exact same culture conditions and incubation times, to avoid differences in ^{15}N-labeled metabolites contents. If different amounts of cells are used to record the two spectra, it is necessary to multiply the global intensity of the spectrum B accordingly before subtraction. Adapted from ref [238]. Copyright 2016 Springer Nature.

Second, in-cell NMR samples produced using *in situ* recombinant production generate spectra with important cellular background signals. It is possible to remove this background by subtracting NMR spectra recorded with mock-transfected cells grown in the same culture medium, and supplemented with the same labeled amino acids (Figure 8). This strategy has been used efficiently by Banci, Ito and their colleagues through the last 10 years.[226,227,237,238,270]

Third, it is important to notice that the evoked recombinant expression systems do not generate homogeneous populations: i) mammalian cells are not all transfected equally, and individual insect cells are not infected by the baculovirus at the same time point; ii) the genetic and phenotypic profiles of these cells is heterogeneous. Hence, some cells can express high quantities of the heterologous protein, while some other do not produce it at all. The homogeneity of the subcellular localization represents a supplementary layer of potential complexity. Most of the in-cell NMR studies using *in situ* expression have reported average intracellular protein concentrations, notably using semi-quantitative western-blots and NMR signal intensities. However, information from flow cytometry and/or fluorescence microscopy is scarce in these reports, to the best of our knowledge. It would be necessary to evaluate the homogeneity of the analyzed cell population, a key aspect for data interpretation. The studies using cellular delivery (see below) provide this information in most cases, probably because this latter approach is more controversial and calls for multiple checks.

2.2.6. Protein/nucleic acid delivery in cells: a brief overview

Instead of the "*in situ* strategy", a second approach consists in delivering in cells an isotope labeled protein/nucleic acid that was previously produced and purified, either recombinantly or synthetically. Several methods have been proposed. These are naturally different depending on the type of cells. We discuss them extensively in the following subchapter, but we delineate a brief landscape first.

Microinjection of millimeter large single-cell oocytes has been the first technique employed to translocate isotope-labeled proteins in the cytosol of live cells. In-cell NMR in bacterial or mammalian cells requires at least a few millions of them: microinjecting these one by one is no longer an option. Other delivery methods have been explored, focusing mostly on cultured mammalian cells. Three of them have been successfully applied, relying on the use of cell-penetrating peptides (CPPs), pore-forming toxins (PFTs) or electroporation. Here again, detailed protocols exist in the literature.[134,271–273] Every approach has its intrinsic capacities and limits. We attempted to draw a comparative analysis below.

It is clear that the efficiency of all these methods vary greatly depending on the protein to internalize. PFTs represent probably an easier and less demanding framework than CPPs and electroporation; as a price to pay, PFTs have also been associated with increased cell instability and leakage until now. PFT-treatment and electroporation have in common to be very demanding in terms of starting material: they require some 10^8 cells and milliliters of cargo at millimolar concentrations to produce an in-cell NMR sample. They ensure a direct delivery without any necessity of endosomal escape, which is a severe obstacle to face with CPPs. Technical advances have been proposed in the recent years for all these methods, which would probably improve the situation.

We shall mention the recent use of heat shock to translocate proteins carrying EPR-dedicated spin labels.[274] Using this approach, proteins of various sizes were internalized, starting with extracellular concentrations of 1 mM and ending with delivered intracellular concentrations between 1 and 7 μM in *E. coli*, and between 0.5 and 1.5 μM in *Pichia pastoris*. Other recent delivery methods might be explored, which take advantage of the many developments in microfluidics: among others, rapid cell squeezing and swelling through microfluidic constriction, possibly combined with microscale electroporation, appear to be sufficiently robust and high-throughput methods for NMR purposes.[275,276]

2.2.7. Protein/nucleic acid delivery in large oocytes

Purified protein/nucleic acid can be microinjected in millimeter large single-cell oocytes like those of *Xenopus laevis* or *Danio rerio*. The latter ones have been utilized only very recently.[277,278] We focus on the better established *X. laevis* oocytes below. These are quite appealing systems for in-cell NMR studies: i) only a ~170 of them are necessary to fill a 5 mm NMR tube; ii) they can be manipulated and microinjected cell by cell in about one hour; iii) they do not require a specific wet-lab, even though a source of oocytes has to be found (commercial companies sell them at ~1-2 $ a piece); iv) they can stand the injection of about 20-50 nL of high-concentration material, which generates ~20-50-fold intracellular dilutions; v) they can survive and maintain their integrity about 18 hours in the NMR tube without any medium replenishment; vi) they are very-well known model systems in biology.[279] This strategy has been used for in-cell NMR investigations of both proteins[280–286,192,287,288,197,196] and nucleic acids.[289–294] We encourage the reader to refer to the published, detailed protocols.[134,146,295–297] We will not present them extensively once again, but we must give a few words of caution: i) because of the cylindrical shape of NMR tubes, hundreds of oocytes must be stacked on each other and thus require high-density outer solvent (e.g. a buffer containing 20% w/w Ficoll) to not crush within a few hours;[283] ii) their cell-cycle stage is well-defined, which comes with certain advantages when studying cell-signaling and post-translational modifications,[282–285,298] but it leaves little room for modifications of the cellular conditions; iii) their high-lipid yolk compartment is problematic for molecules that would preferentially bind to lipids; iv) they generate important peak broadening due to sample inhomogeneities (settled oocytes leave large intercellular spaces, and the intracellular compartments are also diverse enough to present distinct magnetic susceptibilities);[286 62] v) their cellular organization is clearly not that of more standard prokaryotic or eukaryotic cells. Oocyte cytosolic extracts represent a good proxy in many regards, e.g. to test the cellular stability, the NMR detectability or even to investigate post-translational modifications (PTMs).[284,298,299] Extracts are also easier to handle, allow a better control of concentrations, of the time origin for time-resolved NMR monitoring, and better NMR resolution.[286,134,294,179,297] Hence, while they enabled many seminal contributions in the field, oocytes tend to lose their attractivity in the very last years.

2.2.8. Protein/nucleic acid delivery using cell-penetrating peptides

First, let us focus on cell-penetrating peptides (CPPs). These represent a scientific field of its own, focusing mostly on eukaryotic cells,[300–303] although some efforts are also being dedicated to bacterial cells.[304] Their expected role is to enable

the translocation of a cargo protein to which they are linked. CPPs can also help in nucleic acid transduction,[305] which we will not discuss because it has not been used for in-cell NMR purposes, to the best of our knowledge. CPPs penetrate cells via multiple, simultaneous mechanisms including energy-dependent endocytosis and micropinocytosis and energy-independent passive diffusion and direct translocation, which are dominant at low (<5 μM) and high (> 10 μM) CPP concentrations, respectively.[300–303,306] This means that a good proportion of the incorporated CPP-cargo population goes to endosomes, from which it is difficult to escape (Figure 9).[307]

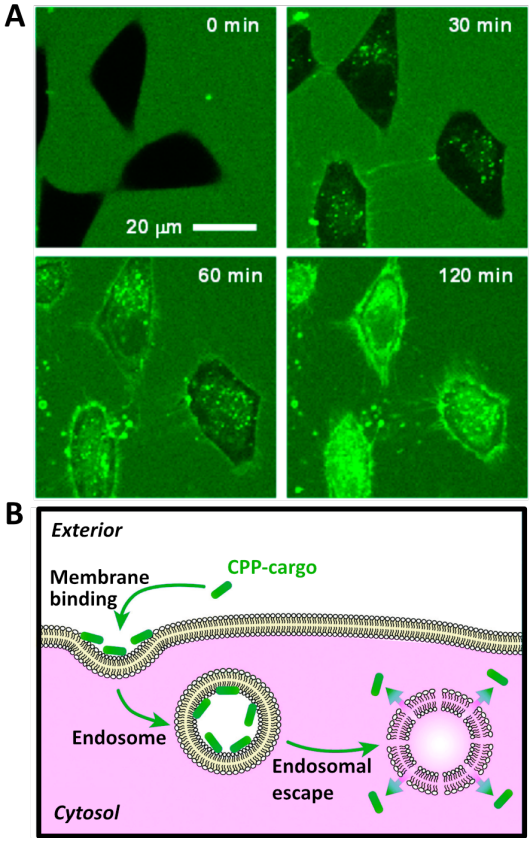


Figure 9: Cell-penetrating peptides penetrate cells via a combination of multiple mechanisms, among which endocytosis is dominant in most cases. **A)** Live-cell fluorescence microscopy of HeLa cells exposed to a FITC-labeled CPP (in green) showing high endosomal escape properties, adapted from ref [308], copyright 2016 American Chemical Society; **B)** Cartoon representation of CPP-cargo penetration in cells via endocytosis; this energy-dependent process involves a number of endogenous proteins, which are not represented here; endosomal escape is variable and can be improved by a number of strategies, but is hardly 100% efficient; adapted from ref [302]. Copyright 2018 John Wiley and Sons.

Optimizing endosomal escape efficiency is useful for drug-delivery, and absolutely necessary for in-cell NMR: if a significant proportion of CPP-cargo molecules remains in endosomes and the rest is released in the cytosol, how do we analyze the resulting average NMR signal? The pitfall here comes from the fact that fixation methods for immunofluorescence tend to permeabilize membranes and provoke the cytosolic redistribution of CPP-cargos trapped in endosomes![301,303,307,309] At the opposite, fluorescent markers can drive CPP-cargo to unwanted subcompartments. As far as in-cell NMR is concerned, another requirement is to cleave CPPs from their cargos. Concretely, Shirakawa and colleagues adopted two strategies: i) binding their CPP and their cargo via a disulfide bond, which got reduced once their CPP-cargo chimera was released in the cytosol; ii) designing a CPP-ubiquitin-cargo chimera, which got cleaved at the ubiquitin C-terminal by undetermined deubiquitinating enzymes ubiquitin.[310–312] They used the popular CPP sequence from the HIVirus-derived TAT peptide. One should notice that Shirakawa and colleagues needed to repeat four exposure+recovery steps (10+40 minutes), cells being incubated with the CPP-cargo constructs at 250 μM during the exposure phase. The obtained intracellular concentrations of the proteins of interest varied between 3 and 100 μM, consistently with the known dependence on the global charge of the CPP-cargo.[310–312] Importantly, they benefited from the help of CPP-expert colleagues, who discovered the synergistic effects of pyrenebutyrate in enhancing direct translocation of CPPs.[306,309,313] As confessed by Shirakawa and colleagues, the process of designing and producing the CPP-cargo is tedious.[314] To the best of our knowledge, it was not adopted by other labs. In the meantime, the field of CPPs raised considerable hopes for therapeutics and has grown consequently. We guess that some of the latest improvements would find applications in the field of in-cell NMR: cultured cells incubated with 1 μM CPP-cargo can internalize it at a final concentration of ~50 μM with satisfying endosomal escape.[315,316] However, the capacity of CPP approaches to deliver the CPP-cargo population with a good synchronicity is not warrantied, although advisable for NMR purposes.

2.2.9. Protein/nucleic acid delivery using pore-forming toxins.

As a second favorite approach, NMR spectroscopists have used pore-forming toxins, more precisely the bacterial Streptolysin O (SLO), to deliver proteins or nucleic acids in mammalian cells. SLO belongs to a family of cholesterol-dependent cytolysins (CDC), which oligomerize to form 20 to 50 nm large pores depending on the CDC and the number of CDC monomers incorporated in the final assembly (Figure 10).[317,318] These pores are removed within minutes upon buffer supplementation with Ca^{2+}: the native intracellular concentration of free [Ca^{2+}] is ~100 nM, and a slight Ca^{2+} influx provokes membrane repair by endocytosis or exocytosis (budding and shedding) of membranes regions surrounding pores.[319,320] However, a number of cells do not reseal properly or fast enough, and intracellular [Ca^{2+}]> 10 μM drives them to apoptosis.[321] SLO permeabilization can also provoke, among others, K+ and ATP effluxes, or swelling, which trigger cell death too.[275,322] Hence, special care must be paid to the buffer composition during SLO-supported delivery. Endocytosis of CDC pores provokes itself the release of endosomes proteases and Ca^{2+} in the cytoplasm.[322] Altogether, resealing upon Ca^{2+} exposure is never uniform in a population of cells, many of them experiencing oscillation or prolonged plateaus of intracellular [Ca^{2+}] leading to delayed cell death and burst.[323] SLO-supported delivery is thus often linked to long-term cellular instability and leakage, as reported by the groups of Shimada or Katahira.[324,325]

To improve the sample quality, Shimada, Nishida and colleagues i) use centrifugation in presence of Percoll to remove dead cells, and ii) evacuate the leaking material using a constant flux of fresh medium in the NMR spectrometer, which they apply on immobilized gel-encapsulated cells.[267,326,327]

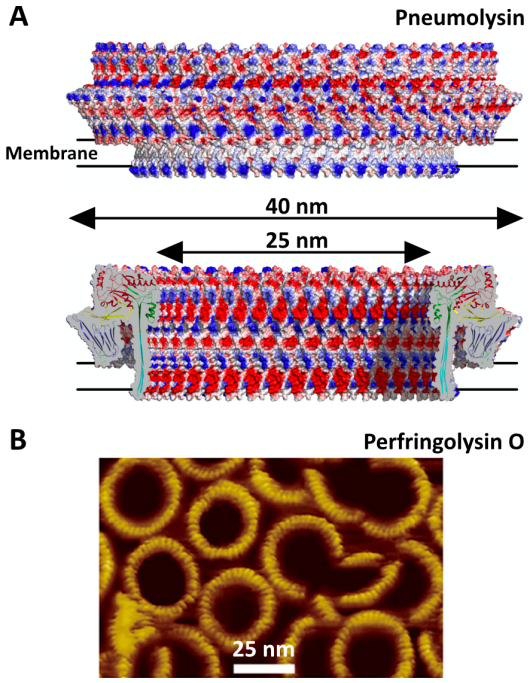


Figure 10: Structural characterization of cholesterol-dependent cytolysins (CDCs), the pore-forming toxin family of Streptolysin O. **A)** CryoEM structure of the pore complex formed by pneumolysin from *Streptococcus pneumoniae*; the electrostatic surface shows the even distribution of positive (blue) and negative (red) patches on the inner surface of the pore complex; adapted from ref [318], copyright 2017 from van Pee et al. under the terms of a Creative Commons CC BY license. **B)** Atomic-Force Microscopy image of the perfringolysin O from Clostridium perfringens; adapted from ref [275], copyright 2018 American Chemical Society.

They reported ~20-30% cell death after 7 to 15 hours of NMR measurements though.[267,326,327] They reported also delivery efficiencies between 60 and 85% of cells incorporating the protein of interest. Finding the best balance between SLO concentration, incubation times, cell death and delivery efficiency requires some tests naturally. To give an idea of the protocols and their yields, Shimada and colleagues incubated cells 30-45 minutes in presence of SLO and of their proteins of interest at 1-2 mM concentrations, which yielded intracellular concentrations varying between ~50 and 180 µM depending on the protein.[267,324,326] Katahira and colleagues used similar protocols for translocating nucleic acids but reported lower final intracellular concentrations of ~5-20 µM, associated to 50% delivery efficiency among the 60% surviving cells.[325,328] Xu and colleagues estimated that they inserted their fluorinated DNA quadruplex at an intracellular concentration of 150 µM, starting from an incubation in presence of DNA at 3 mM.[329] An advantage of SLO is their ease of use, as compared to the other delivery methods. However, one can notice they require high concentrations (and thus high amounts) of the protein to deliver, which comes with a number of possible troubles. Like for CPPs, the use of SLOs has not been transferred to many laboratories, to the best of our knowledge. It is still attractive in our opinion, but it would require the addition of a resting step of a few hours between resealing and the NMR measurements, to remove cells that are meant to collapse even though they do not immediately.

2.2.10. Protein/nucleic acid delivery using electroporation

A third method gained interest in the last years: electroporation. This concept is well-known and used since the 1980's by biologists and biochemists, who need to transfect cells with coding DNA to trigger recombinant expression.[275,330] The fact that electroporation is also capable of provoking the translocation of all sorts of molecules, including proteins. The group of Selenko took advantage of it a few years ago,[176,331] and the method seems to be more and more popular in the community, both for proteins and nucleic acids delivery. Electroporation methods for plasma membrane permeabilization rely on the application of voltage pulses of about 1 kV/cm during a few hundreds of microseconds in 1-2 mm wide cuvettes (Figure 11). They provoke the appearance of membranes pores that can have broad distributions of diameter, between 1 and 50 nm, which reseal spontaneously within seconds to minutes. The window of useful electric field values is limited, because cell death occurs soon for values above 3 kV/cm. The mechanisms of pore formation have been studied, and the literature shows that many parameters (cell type, pulses duration, strength and number, extracellular medium, actin cytoskeleton status, temperature, cargo size, ...)[332–337,275] and a broad range of timescales (pores reseal within seconds, but cells remain permeable to small molecules during hours after the pulses)[336,338–340] must be considered.

We can keep in mind a few key aspects of electroporation-assisted delivery: i) threshold values (in the order of magnitude of 1 kV/cm) and non-linear behaviors exist for pore formation, pore diameter populations and cell survival in reaction to electric field intensity;[275,341] ii) 10 to 1000 ns long pulses generate small pores (< 1 nm), while millisecond long pulses favor larger pores and macromolecules penetration in cells within the next few seconds;[333,336,342] iii) extracellular proteins and small DNA/RNA (below 20 bp) diffuse inside the cell while pores are open, while large DNA associate with the plasma membrane and are integrated via endocytosis;[343] iv) electroporation-assisted delivery increases with temperature;[275,344,337] v) cell permeabilization and intracellular content efflux correlate with long-term (hours) cell death, most pores resealing immediately (50 ms to 10 s for cells) but the proper barrier functions towards ions and small molecules taking up to 20 hours;[275,337,338,342,345] vi) electroporation provokes the appearance of reactive oxygen species and notably lipid peroxidation, which might be correlated to long-term membrane permeabilization;[275,336,342,346] vii) mitochondrial and nuclear ultrastructures, cell homeostasis, and ATP levels take two to five hours to recover their native levels.[275,347] These considerations show that electroporation calls for a case-by-case optimization of the extracellular buffer (e.g. conductivity, ATP supplementation, …),[275,348,349] of cell density, of the number, duration and strength of pulses, of the resting conditions. A balance shall always be foundbetween delivery efficiency and cell survival rate. Importantly, optimal delivery can be achieved by applying tailored sequences of one intense, short and one weaker but longer pulse consecutively, the second leaving pores open with a reduced power deposition.[350,351] We should also add some prosaic observations: the application of electric pulses provokes temperature jumps in the sample (up to 10-20 degrees depending on the extracellular medium)[352,353] and the apparition of foam (partly due to water electrolysis),[354]

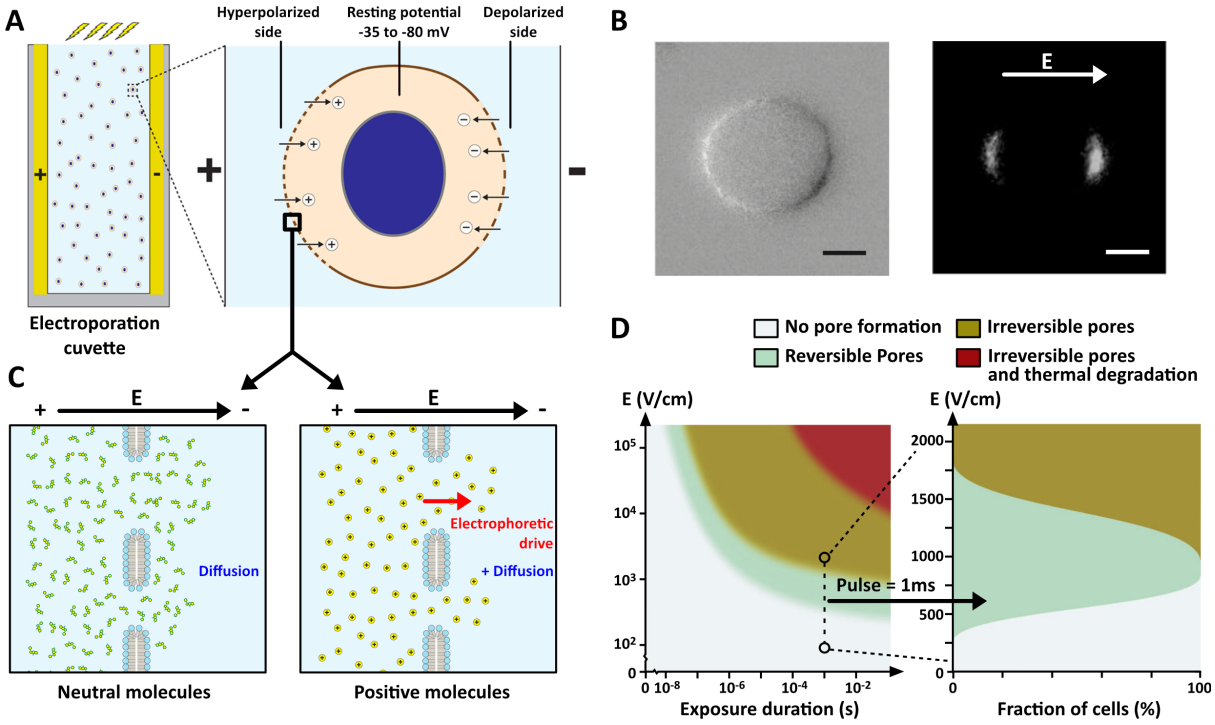


Figure 11: A) Cartoon representation of the conventional electroporation approach, where cells in suspension are placed in an electroporation cuvette and exposed to an electric field; this provokes pore formation that are unequally distributed on cell membrane, whose size and resealing time depend on the strength and duration of the pulse (and on the type of cells of course, typically on the ratio between cell diameter and membrane thickness); adapted from ref [275], copyright 2018 American Chemical Society; **B)** Microscope view of a spherical CHO cell exposed (left) and transport of propidium iodide 200 ms after applying an electroporation pulse of 1.5 ms/650 V/cm (bar= 5 µm); adapted from ref [355], copyright 2010 Springer Nature; **C)** Cartoon representation of cell penetration by 1-10 nm large molecules, adapted from Stewart et al. Chem Rev 2018 [275]; **D)** Schematic diagrams of the effects of electroporation pulses, depending and their strength and length (left), and of the distribution of these effects among a cell population for a 1 ms long pulse (right), adapted from ref [356], copyright 2015 Elsevier.

which are both unfavorable for cell survival and protein stability. Moreover, local pH changes in the close vicinity of the electrodes (up to 5 units of pH) have also harmful effects.[354] Altogether, optimizing these multiple parameters becomes soon time-consuming. Moreover, many commercialized electroporation devices do not communicate their proprietary pulse sequences, which makes the optimization difficult to rationalize. These "black box electroporators" have been used by the groups of Selenko,[176,331,357,358] followed by those of Shekhtman, Riek and Hiller, Baldus, and Petzold.[359,360,71,361] Better controlled systems have been used by the groups of Ito,[362,363] followed by those of Trantirek,[294,364,365] Oliveberg and Danielsson,[366] and Liu[367]: those working on proteins applied a first intense pulse (~100 V, 15 ms) followed by milder but longer pulses (~20 V, 50 ms*2); Trantirek and colleagues, working on nucleic acids, applied a stronger first pulse (1000 V, 100 µs) and a second milder one after a delay of 5 s (350 V, 30 ms). All these groups used low-potassium, high-magnesium electroporation media, often supplemented with ATP and glutathione. Electroporation was executed on cells at densities varying between 10 and 20 million cells per mL, mixed with the protein or nucleic acids of interest at millimolar concentrations (from 0.35 to 2.5 mM, depending on the study). The common used cell lines were HeLa, HEK293 or A2780, but electroporation proved to work efficiently on other cell lines considered to be more sensitive (RCSN-3, SH-SY5Y).[176,331,357] Cells were allowed to rest 3 to 5 hours on plates in most cases, which permit also to discard cells that do not attach and float. Overall, these protocols yield intracellular concentrations of proteins between 3 and 50 µM and of nucleic acids between 10 and 20 µM.

An important concern has been a recurrent subject of informal discussions through the years: electroporation might unfold the protein of interest. A number of in-cell studies were soon considered to come to a dead end after observing awkward in-cell spectra showing broad, poorly informative NMR signals in the regions where unfolded peptides resonate. Many causes can lead to such spectra. One of them could be unfolding provoked by electroporation. Theoretically, the orders of magnitudes do not fit: according to a number of simulations, protein unfolding by an electric field requires intensities of 10^6-10^7 V/cm,[368–375] much higher than the ~10^3 V/cm commonly used for electroporation delivery. Quantum mechanics calculations predict that a polyalanine helix would be disrupted from 2.5x10^6 V/cm. Applying 10^6 V/cm on a protein crystal results in faint structural changes.[376] However, experimentally, lower electric fields can have deleterious consequences on protein structure or activity, either fields of 5-50x10^3 V/cm applied during a few microseconds,[377–380,108,381] or 50-500 V/cm applied through minutes to hours.[382,383] Frictional forces due to electrophoretic motion of the corresponding purified proteins might be the cause of these observations.[382] These scenarios do not correspond to the common pulsing protocols used for electroporation delivery. Hence, we tend to think that electro-

poration-provoked unfolding might be an exception but not a rule. Interestingly, antibodies[384–388] and kinetochore components[389] have been successfully "electroporated" in cells, where they displayed functional behaviors. Moreover, in-cell EPR studies did not report any protein unfolding upon electroporation yet.[390–394] (N.B.: It is always possible to verify that a protein kept its proper fold upon electroporation in absence of cells. However, the result will not rule out the possibility that destabilized conformations might refold correctly *in vitro*, while they might be trapped upon binding to cellular constituents once they have penetrated cells.)

2.2.11. Flow-probe bioreactors and gel encapsulation.

A recurrent concern in in-cell NMR studies is the health status of cells in the NMR tube. A high cell density is often necessary to reach sufficient S/N, which leads to rapid nutrient exhaustion in absence of replenishment of the culture medium. Among others, intracellular ATP exhaustion occurs in a couple of hours in mammalian cell pellets (Figure 12), which had to be used to carry out the pioneering studies.[267–269,395] Some studies reported experimental evidences of the impact of the cellular metabolic state on protein conformations and cellular interactions.[160,206,363] It would be advisable to set up flow-probe bioreactors allowing steady conditions: these would ensure constant fluxes of fresh nutrients and withdrawal of metabolic end-products.

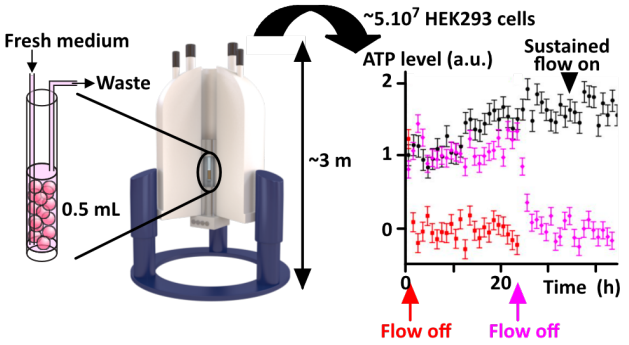


Figure 12: ***Left:*** Schematic presentation of a flow-probe bioreactor, where a constant flow of fresh medium (50-100 µL/min) is maintained in the NMR tube during the NMR acquisition; ***Right:*** ATP levels (as measured by ^{31}P-NMR) in the NMR tube through extended periods of time in presence of a constant medium flow (black), a flow interrupted after 24 hours (magenta), or in absence of any medium replenishment (red), the sample being composed of ~3.10^7 HEK293T cells in agarose threads; adapted from ref [269]. Copyright 2020 American Chemical Society.

In a recent inventory, we counted about 200 publications reporting the use of NMR-compatible flow-probe bioreactors.[62] Most of them were actually proposed before the year 2000, in times where large diameter NMR probes were used to monitor organs and cells metabolism. We refer the reader to our detailed inventory spanning this wider range of in-cell NMR studies.[62] Sadly enough, these flow-probe systems were developed in-house and were not transferred from lab to lab. The modern 3-5 mm narrow (cryo)probes offer much better sensitivity, but less flexibility to plug additional accessories. In-house and commercial devices were adapted to these up-to-date NMR probes, with the purpose of maintaining cells healthy for in-cell structural biology studies. These systems require most often to encapsulated cells in gel matrices to avoid them to drain off with the outgoing medium. Only few structural biology NMR groups had time and manpower to establish them though.

In 2010, Pielak and colleagues proposed a flow-probe bioreactor, which permitted alternated spikes of fresh nutrients, cells settling, and NMR measurements of αSynuclein overexpressed in *E. coli*.[155] Their system was adapted to a 8 mm diameter tube and probe, an unconventional system that does not favor sensitivity and magnetic field homogeneity. They encapsulated cells in 1 mm large alginate beads, which apparently affected further the magnetic field homogeneity and yielded broader NMR peaks.

In 2013, Shimada, Nishida and colleagues reported the use of a home-made system adapted to commercial 5 mm cryoprobes.[267] A syringe pump delivered fresh DMEM medium at a flow rate of 50 µL/min, through a 1 mm polymer tubing inserted in the 5 mm NMR tube. DMEM was degassed and buffered with HEPES instead of the more common HCO_3^- to avoid air bubbles. The excess liquid was removed via another polymer tube connected to an aspirator. 3.10^7 HeLa S3 mammalian cells were sequestered in threads of matrix "gelifying" above 25 ºC (a thermoreversible hydrogel so-called Mebiol gel), and maintained viable at 37 ºC: 80 % viability was reported after 15 hours, instead of 20% in a cell pellet without medium replenishment.[267] The deuterium lock was ensured by the D_2O used to dissolve the hydrogel. The authors reemployed the same system in two recent studies.[326,327] Interestingly, they showed that they could control the redox conditions in the flow-probe, adding tert-butyl hydroperoxide in DMEM to generate oxidative stress.[326] Katahira used the same flow-probe system to record NMR spectra of DNA oligonucleotides in cells embedded in agarose threads.[328]

Shekthman and colleagues proposed another home-made set up, which uses a gravity siphon to supply fresh medium at a flow ratee of 100 µL/min in the NMR tube via a 0.5 mm PTFE tubing, in which 50 µm holes were drilled at its extremity.[395] They improved recently the medium delivery by using an ultrahigh molecular weight microporous polyethylene diffuser tubing and a peristaltic pump to control the medium waste output.[396] They initially used agarose threads to maintain *E. coli* and 4.10^7 HeLa cells in a 5 mm NMR tube at 25 ºC, allowing 99% viability of HeLa cells after 24 hours. Agarose threads formed in 100% D_2O were added in the NMR tube enable the lock. In their second report, they used 1 mm large alginate beads to encapsulated *E. coli* cells, because these beads can expand and permit cell growth.[396]

From 2019, Banci, Luchinat and colleagues have adopted a commercial flow-probe bioreactor designed to deliver an input flow via an inlet polymer tubing and withdraw the excess liquid via a coaxial outlet tubing. Initially, they plugged a cylindrical 1 MDa cut-off dialysis membrane to this coaxial system and plugged it into the 5 mm NMR tube. They used a flow rate of 50 µL/min in a mammalian cell suspension hold in 30% Percoll to avoid sedimentation.[268] A glass capillary containing 100% D_2O was inserted for the deuterium lock. This system did not ensure a homogeneous cell survival (Luchinat, personal communication). In their more recent publications, the authors encapsulated 3.10^7 cells in agarose threads and delivered fresh medium (supplemented with 2% D_2O) in the NMR tube without capping the coaxial inlet/outlet tubing with a dialysis membrane.[269,397]

Interestingly, Ito and colleagues also used a flow-probe to maintain insect cells viable in a NMR tube through extended periods of times, but they did not have to encapsulate these cells: they delivered fresh medium at flow rate of 2 mL/hour, which simply kept cells in suspension.[227] Adherent cells have still to be entrapped in hydrogels. In this regard, methyl cellulose has been shown to have improved impact on cell viability in two recent live-cell NMR studies.[398,399] An important aspect here is to avoid to entrap any air bubble in the hydrogel: this would provoke important magnetic susceptibility inhomogeneities, which drastically affect NMR linewidth.[62]

2.3. Solution NMR techniques.

2.3.1. Chemical shifts report for structure, post-translational modifications and interactions

We shall first recall that NMR chemical shifts are exquisite reporters of the chemical environment, i.e. of the structure, of the potential chemical modifications, and of the intra- or inter-molecular interactions.[400–410] Hence, chemical shift perturbations reveal structural or chemical changes in the vicinity of the observed nuclei. These chemical shifts are necessarily recorded and correspond to the primary information extracted from NMR spectra.

Spectroscopists use somehow standard NMR pulse sequences for in-cell NMR studies, but they have to pay attention to the deleterious effects of cellular viscosity and of multiple interactions with cellular components: these provoke slower tumbling and chemical exchange, hence faster T2 relaxation and lower NMR signal (see chapter 2.1.3.). As a rough simplification, a trade-off is often to be found between the intensity and the informative capacity of in-cell NMR signals. One-dimensional ^{19}F-NMR spectra generate sometimes a single, broad signal, which can be more useful than a 2D ^{1}H-^{15}N spectrum where hundreds of residue-specific peaks are broadened beyond detection in cellular conditions. We explain in the following which pulse sequences have been used to carry out in-cell structural biology by solution NMR, and we give some elements to understand the rationale behind.

2.3.2. Which pulse sequences to use for 2D ^{1}H-^{15}N correlation spectra?

^{15}N-editing coupled to ^{1}H-detection has been the workhorse of in-cell structural biology by NMR, notably because it can generate spectra with low cellular background in comparison to ^{13}C-edited spectra.[146,147,226,227] This strategy is often used to acquire well-dispersed two-dimensional spectra, where every ^{1}H-^{15}N NMR correlation peak correspond to one amide function, hence eventually to one amino acid. A number of pulse sequences can execute such ^{1}H-^{15}N correlation spectroscopy, which we can split in two families, namely the HMQC and HSQC schemes (Heteronuclear Multiple/Single Quantum Correlation). The fine description of these sequences is beyond the scope of this review, and we refer the reader to dedicated publications.[411–417] For the following discussion, we must explain that: i) these pulse sequences have been modified to speed up the scanning pace (by enhancing the T1 relaxation), and thus to increase the signal-to-noise ratio per unit of time (S/N-PUT), and named ^{1}H-^{15}N SOFAST-HMQC[411,412] and BEST-HSQC[416] (or LHSQC)[414]; ii) versions of the HSQC pulse sequences exist that are adapted to large proteins, called ^{1}H-^{15}N HSQC-TROSY[413,417] and the combination BEST-TROSY (or LTROSY).[416,418]

Spectroscopists have to decide between these different pulse sequences, which all have advantages and drawbacks, particularly in the cellular context. Let us give a rough summary: i) in the case of a fast-tumbling protein (~5-10 kDa, or IDPs), the S/N-PUT is as follows ^{1}H-^{15}N SOFAST-HMQC > BEST-HSQC > HSQC > BEST-TROSY > HSQC-TROSY; ii) in terms of peak linewidth, hence of potential resolution, the order is strictly the opposite; iii) the relative S/N-PUT of pulse sequences is progressively inverted for proteins of growing size (i.e. slower tumbling-time), so that only ^{1}H-^{15}N TROSY pulse sequences can help to detect proteins above 100 kDa for example.

Which ^{15}N-editing pulse sequences were thus used for in-cell studies? Dötsch and colleagues pioneered the field using ^{1}H-^{15}N HSQC, at a time where BEST updates had not been designed yet.[143,146,147] This pulse sequence was good enough for detecting the small model protein NmerA (7 kDa) in *E. coli*, and even better for another small model protein, GB1 (6 kDa). Indeed, GB1 experienced only limited interactions with cellular components, so that its tumbling was ~8-fold slower (i.e. like a protein of ~50-60 kDa *in vitro*) in *E. coli* than *in vitro*, which yields still relatively favorable NMR relaxation properties and signal intensities.[173] Two-dimensional ^{1}H-^{15}N HSQC was thus used through the 2000's by various group to detect well-behaved folded proteins like GB1 and some IDPs in *E. coli*, *X. laevis* oocytes, yeast, insect and mammalian cells.[149,153,165–168,182–184,217,226,282,283,324,419] However, it was soon recognized that some small proteins did not yield any detectable signal in cells, the best example being ubiquitin.[148,153,156,170,173,280,420] Multiple interactions with cellular components were responsible of this peak disappearance: these provoke slower tumbling but also chemical exchange relaxation, which, altogether, cannot be solved by the sole use of TROSY pulse sequences.[146,148,189,285,421] The physiological state of bacterial cells plays a role in presenting variable quantities of sticking RNAs or chaperones and in producing eventually misfolded populations of the protein of interest: here again, TROSY sequences find their limits and cannot help to detect any useful signal.[206,359] Small folded (10-25 kDa), promiscuous-interacting proteins, like ubiquitin or thioredoxin, appear to tumble as if they were ~1 Mda large in *E. coli*.[359] Hence, only deuteration of these latter proteins is likely to make them NMR detectable, in association with dedicated TROSY experiments.[190,286,359,395,422]

In 2009, Shirakawa and colleagues were the first to use the ^{1}H-^{15}N SOFAST-HMQC to acquire 2D NMR spectra of 6-12 kDa large proteins in mammalian cells.[310] Because of its superior S/N-PUT, Banci and colleagues followed this ^{1}H-^{15}N SOFAST-HMQC approach for studying a 7 kDa large protein and the 32 kDa large SOD1 homodimer in *E. coli*.[171,423] They used it later for characterizing proteins of various weights in mammalian cells, including the SOD1 homodimer.[172,237,238,255,256,259,260,269] Then, the ^{1}H-^{15}N SOFAST-HMQC pulse sequence became the first choice of a number of groups for characterizing folded proteins in mammalian cells in the 2010's.[285,311,312,331,357,358,360,362,363,366,396,424,425] The mammalian cytosol is less crowded and viscous than the bacterial one, so that the studied small proteins (< 20 kDa) have generally

intracellular tumbling times about twice slower than in their purified, dilute form. Hence, mammalian cells do often not provoke the great intensity losses that 2D ^{1}H-^{15}N SOFAST-HMQC suffers with large proteins (> 50-60 kDa). At the opposite, some groups kept using the ^{1}H-^{15}N HSQC when working in *E. coli* or frog/fish oocytes, where higher viscosity provokes faster relaxation that is more deleterious to 2D ^{1}H-^{15}N SOFAST-HMQC than to ^{1}H-^{15}N HSQC acquisition.[180,185–187,201,278,287,288,426,427]

We must mention the specific case of IDPs: although the ^{1}H-^{15}N SOFAST-HMQC approach was initially tailored for globular proteins, it can yield 4-20-fold S/N enhancements for IDPs' detection *in vitro*, notably because of a favorable side-effect of the water-amide proton exchange.[428] It has thus been used for in-cell NMR studies of IDPs.[176,177,331,357,360] However, here again, the higher intracellular viscosity and the multiple transient interactions provoke a fast relaxation in the ^{15}N-dimension and thus a rather poor resolution, even at very high field.[270] Unfortunately, IDPs generate poorly dispersed NMR spectra, which calls for highly resolved spectra if residue-specific information is expected. Hence, a number of groups preferred to acquire 2D ^{1}H-^{15}N HSQC spectra of IDPs in cells when their concentration was high enough,[284,360,367,426] or even better to use BEST-HSQC (LHSQC).[176]

To the best of our knowledge, nucleic acids have not been studied in cells using ^{1}H-^{15}N NMR correlation spectroscopy. Pulse sequences dedicated to nucleic acids exist, which are relatively equivalent to those developed for proteins.[418]

To finish with ^{1}H-^{15}N pulse sequences for in-cell NMR, we must discuss their field dependency. The NMR spectroscopists know that the performances of TROSY sequences do highly depend on the magnetic field. If we consider an ideal, deuterated ~30 kDa protein (without any conformational or binding exchange), theoretical calculations indicate that ^{1}H-^{15}N non-TROSY and TROSY peak intensities should be ~2-fold and ~4-5 fold larger at 1.2 GHz than at 500 MHz, respectively.[429] For such a protein, TROSY schemes become more sensitive than regular HSQC sequences only at ~1 GHz fields. The efficiency of ^{1}H-^{15}N SOFAST-HMQC and BEST-HSQC (LHSQC) experiments depends on field too: higher fields permit to excite more selectively the HN amide resonances, and thus to maintain the magnetization of water and alkyl protons along +z to a better extent, both populations accelerating the HN T1 relaxation and its S/N-PUT.[412,428] Hence, using the current commercial equipment and pulses, these pulse sequences start to provide their best yields at fields above 700-800 MHz, according to our practical experience. What can we expect from higher magnetic fields for in-cell NMR investigations? The first reports are

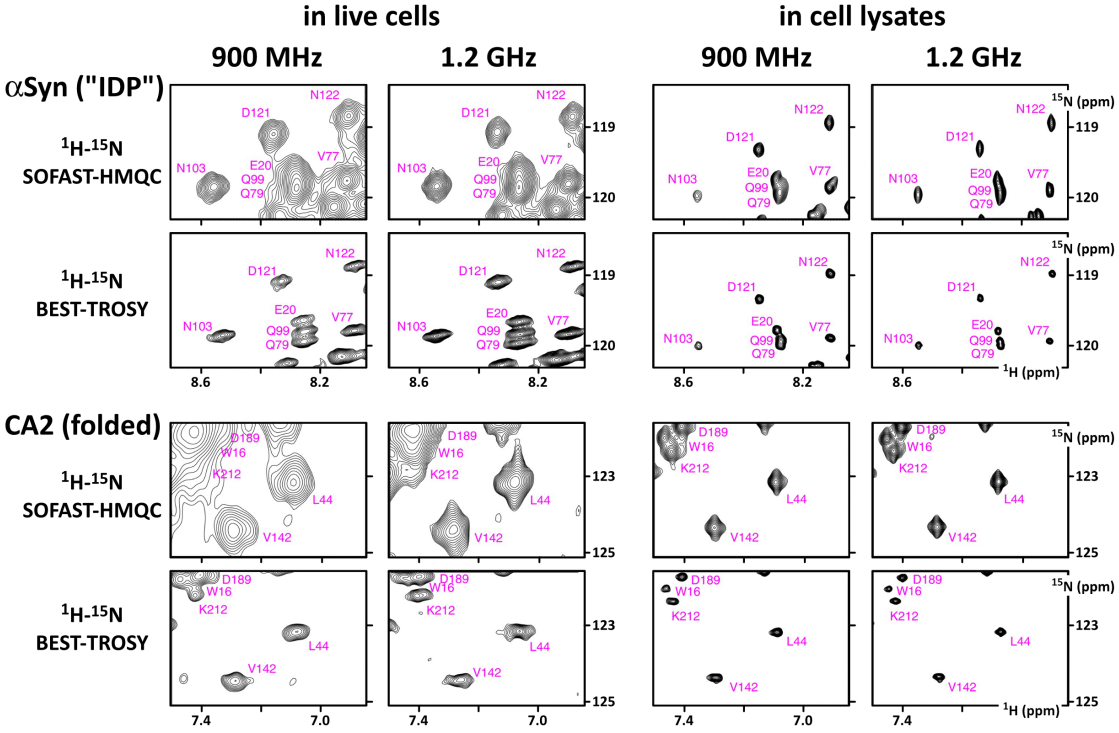


Figure 13: In-cell ^{1}H-^{15}N SOFAST-HMQC and BEST-TROSY spectra of the model disordered protein α-synuclein, and of the folded CA2 (~30 kDa), recorded in HEK293T cells (after transient transfection and *in situ* expression in a medium supplemented with ^{15}N-labeled amino acids).[270] Intracellular concentration of CA2 was ~150 µM, according to previous reports (unknown for α-synuclein).[261] The lowest contour levels were set to 4x the average noise intensity. All SOFAST-HMQC experiments were recorded in ~30 minutes, except that of α-synuclein in live cells in ~1 hour; all BEST-TROSY were recorded in ~1 hour. The evolution times were set to different values depending on the expected relaxation times and linewidths, but were kept constant for every comparative couples 900MHz-1.2GHz. This means more points in the ^{15}N dimensions and compensating adjustments of the number of scans at 1.2 GHz, in order to keep similar acquisition times. While the higher field brings some improvements in lysates, methods remain to be found to make it profitable for in-cell samples. Adapted from ref [270] Copyright 2021 from Luchinat et al. under the terms of a Creative Commons CC BY license.

only being released on the brand-new 1.2 GHz spectrometers. Banci and coworkers showed that S/N-PUT of ^{1}H-^{15}N SOFAST-HMQC and ^{1}H-^{15}N BEST-TROSY experimentswere about 75% more sensitive at 1.2 GHz (3-mm

probe) than at 900 MHz (5-mm probe) for the model IDP α-synuclein in mammalian HEK293 cells (up to ~130% in cellular extracts); the same pulse sequences were approximately as sensitive at both fields for the ~30 kDa enzyme carbonic anhydrase 2 in the same cells (CA2).[270] These numbers come from only 3 selected peaks in the acquired spectra. Moreover, they were biased by the fact that they compared samples in 3-mm diameter tubes, but recorded using 5- and 3-mm diameter probes at 900 MHz and 1.2 GHz, respectively: S/N-PUT at 900 MHz might be increased by ~30% to account for the lower filling factor of the 5 mm probes. The ^{1}H and ^{15}N NMR peak linewidths were 10-20% thinner (in ppm) at 1.2 GHz. Altogether, in-cell solution NMR samples may benefit of "ultra-high" fields only at the margins, knowing that these two model proteins are well-behaved for our purposes, i.e. they do not interact too promiscuously in cells and they do not show µs-ms conformational dynamics. It seems that we observe the typical compensation between the beneficial (a higher starting magnetization) and adverse (faster T2 relaxation, and also slower T1 relaxation) effects of increasing magnetic fields on protein NMR signals. The ultra-high field might be more advantageous for ^{13}C/^{15}N-direct detection.

2.3.3. ^{13}C-edited spectra

As mentioned several times in chapter 2.2., 2D ^{1}H-^{13}C correlation spectra contain more signals from the cellular background than ^{1}H-^{15}N spectra. However, ^{1}H-^{13}C NMR spectroscopy of methyls has advantages that made it instrumental for characterizing large proteins *in vitro*: i) methyls are rather mobile moieties, ii) they carry 3 equivalent protons, iii) dedicated pulse sequences for high molecular weight proteins are relatively straightforward to use, and iv) a number of metabolic precursors and labeling protocols exist to integrate well-defined ^{1}H-^{13}C-methyls in a deuterated background.[85,104] Of note, the SOFAST-methyl-TROSY is the recommended pulse sequence in all situations and its performances are relatively field-independent.[85,430] All these combine to make 2D ^{1}H-^{13}C NMR of methyls a sensitive and convenient approach. However, it generates spectra that are intrinsically less dispersed than 2D ^{1}H-^{15}N spectra. Thus, 2D ^{1}H-^{13}C NMR helps only for folded proteins and not for IDPs. Moreover, the cellular background is quite profuse in the spectral regions of Ala, Leu, Thr and Val.[146,148,267] It prohibits most probably studies in *X. laevis*, due to their high lipid content.[286] In 2004, Dötsch and colleagues showed that [methyl-^{13}C]-methionine labeling could be helpful to detect proteins in *E. coli*. They were using a ^{1}H-^{13}C HSQC pulse sequence, which is nowadays obsolete (but also used by Banci and colleagues in 2016).[85,238]

The next in-cell ^{1}H-^{13}C spectra came only in 2012 by Waudby and colleagues, who reported methyl-TROSY spectra of [^{13}Cδ]-Ile of the model protein TTHA in *E. coli*; the peaks were rather broad.[174] Shimada and colleagues produced ^{1}H-^{13}C SOFAST-methyl-TROSY of a 9 kDa protein delivered in mammalian cells. The protein was uniformly deuterated except the ^{1}H-^{13}C-labeled methyls on Ile-δ1, Leu-δ1/2 and Val-γ1/2 positions, which permitted to recorded high resolution spectra in cells.[267] They noticed moreover that mammalian cells viability was not affected if maintained in 60% D_2O, which helped to record higher quality spectra. Similar to *in vitro,*[85] protein deuteration appears to be key to obtain good resolution and signal intensity. The same authors recorded good-looking in-cell ^{1}H-^{13}C SOFAST-methyl-TROSY spectra from [u-^{2}H, Ala-$^{13}CH_3$]-thioredoxin,[326] and [u-^{2}H, Ile-(δ1-$^{13}CH_3$)]-RAS (20 kDa). In the first case, they monitored the residue A29 of thioredoxin: it had indeed unusual resonance frequencies, out of the spectral regions where most cellular background signals are found. In the second case, Ile-(δ1-$^{13}CH_3$) labeling provided very satisfying spectra in a region where the cellular background is scarce (Figure 14). Moreover, Ile residues are satisfyingly abundant and dispersed in proteins, and Ile-(δ1-$^{13}CH_3$) provides generally more precise structural information than Met-($^{13}CH_3$). This strategy looks promising. However, ^{1}H-^{13}C SOFAST-methyl-TROSY does not solve all the intracellular, broadening effects. In particular, it failed to detect proteins that experience multiple, transient interactions with cellular components in the more crowded cytosol of *E. coli* cells (even though it appears to perform a bit better than ^{1}H-^{15}N BEST-TROSY experiments).[189]

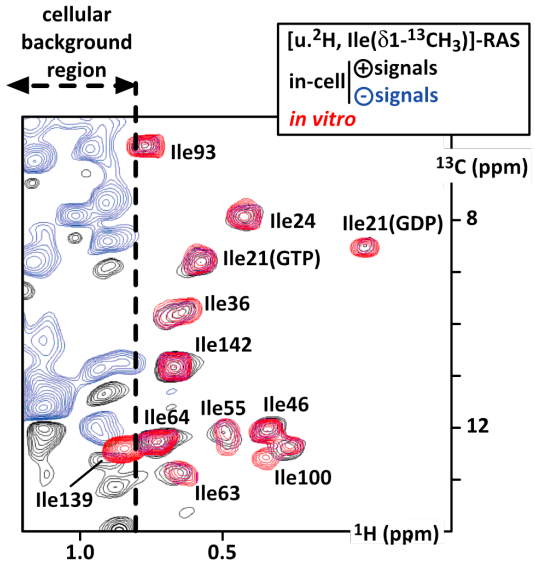


Figure 14: 2D ^{1}H-^{13}C SOFAST-methyl-TROSY spectra of [u-^{2}H, Ile-(δ1-$^{13}CH_3$)]-RAS (20 kDa) in cultured mammalian cells (3.10^7 HeLa S3) (black and blue for positive and negative signals), and in dilute solution (red for positive signals). The acquisition time is 30 minutes for an intracellular concentration at ~50-70 µM. The binding of GTP or GTP to RAS generates two well-separated NMR peaks for Ile21. All Ile signals are observed in cells, except that of Ile139 because it overlaps with cellular background signals. Adapted from ref [327]. Copyright from Zhao et al. nder the terms of a Creative Commons CC BY license https://creativecommons.org/licenses/by/4.0/.

Tridimensional ^{1}H/^{13}C/^{15}N experiments are another common family of ^{13}C-edited spectra, which have been recorded using in-cell NMR samples. These spectra are usually recorded for protein resonance assignments, and eventually for structure determination (see chapter 2.3.5.). Their detailed description is far beyond the scope of this review, and we refer the reader to dedicated publications.[431–433] These experiments are even more sensitive to the adverse effects of cellular interactions and viscosity, because they rely on multiple magnetization transfers during which fast T2 relaxation is extremely deleterious. Moreover, n-dimensional spectra correspond to a geometric increase in data points to record, which translates into days-long acquisition times for many 3D ^{1}H/^{13}C/^{15}N experiments. Sub-sampling and recon-

struction methods have been developed,[433,434] which decrease acquisition times to a few hours. These proved to be instrumental for in-cell samples. Hence, using radial projection and reconstruction, the first in-cell 3D spectra (HNCA in 2 hours, HNCO in 1 hour, and HA(CA)NH in 3 hours) were recorded in 2005 on GB1 in *E. coli*. We have seen previously that the 6 kDa model protein is not too affected by the bacterial cytosol, where it tumbles like a ~50-60 kDa proteins. The authors did not provide the protein concentration, but GB1 overexpression can yield intracellular concentrations about ~250-1500 μM.[169,170] Then Ito and colleagues used non-uniform sampling (NUS) to record 3D spectra of the model protein TTHA1718 at 3-4 mM in *E. coli*.[167,168] According to their relaxation measurements, TTHA1718 was tumbling like a ~40 kDa protein at 37°C in cells. They recorded twelve 3D spectra for backbone- and side-chain assignment as well as distance measurements (see the list in [168]), sampling between 12.5 and 25% of the data points. These spectra could not be all recorded on the same sample, because cells start to leak after 5-6 hours.

Ito and colleagues managed to improve their approach by reconstructing their 3D spectra using Quantitative Maximum Entropy algorithms, which shows improved reliability including for NOESY spectra.[435] This permitted to record the same set of 3D spectra of GB1 at only ~250 μM in *E. coli*, at 22°C, using 140 minutes long acquisition times.[169] Ito and colleagues had previously used this reconstruction method to obtain 3D spectra of GB1 in insect cells, but the lower concentration (~130 μM) permitted only the backbone-assignment and distance measurements. Indeed, the authors failed to obtain exploitable spectra for side-chain assignment, because of the fast T2 relaxation provoked by intracellular viscosity and transient interactions.[226] They fixed this problem by implementing a flow-probe bioreactor supplying fresh medium to the cells, which maintained their viability for 24 hours and more (instead of < 8 hours previously).[227] Hence, they could record the whole set of 3D spectra for five different model proteins (MW ranging from 7 to 19 kDa) in insect cells at 28°C. All these acquisitions were performed on bacterial and insect cells resuspended in 100% D_2O media. The in-cell spectra of GB1 were good enough to assign unambiguously 98% of backbone resonances and more than 70% of side chain resonances. It is important to notice that all these spectra were recorded at 600 MHz, a standard and widespread magnetic field, which is however limited in its resolution capacities. Ito and colleagues used standard 3D pulse sequences, without using band-selective excitation methods (so-called BEST), which might nevertheless significantly improve the S/N-PUT at higher fields.[416]

IDPs have also been studied in cells using ^{13}C-edited experiments. Waudby and colleagues have recorded 3D BEST-HNCO and BEST-HNCOCACB spectra of α-synuclein at ~ 1.7 mM in *E. coli* at 4°C and 700 MHz (~2.5 hours of acquisition time for each spectrum). Interestingly, they used a deconvolution algorithm to compensate the broadening effects of field inhomogeneity, which yielded much improved peak linewidth and resolution. ^{13}C-direct detection NMR spectroscopy was also proposed by Felli and Pierattelli: they showed the feasibility of recording 2D (H-flip)^{13}CO^{15}N and (^{1}H-flip)^{13}Cα^{13}CO spectra on α-synuclein at ~ 0.2 mM in *E. coli* (at 700 MHz, using a carbon-detection optimized probe).[178] Even though less sensitive than ^{1}H-^{15}N HSQC experiments, these ^{13}C-detection experiments provided much improved spectral dispersion. Moreover, the ^{13}Cα^{13}CO experiments can produce spectra of IDPs in all conditions of pH and temperature, which is not completely the case for ^{13}CO^{15}N spectra, and not the case at all for ^{1}H-^{15}N schemes, because of the water-amide proton exchange.[436] Recent ^{13}Cα^{13}CO pulse sequences were proposed, which can produce exploitable 2D spectra of proteins at ~20 μM in ~1 hour.[436,437] ^{13}CO-direct detection approaches are however not likely to be efficient for globular proteins in cells, because of the large chemical shift anisotropy of peptide ^{13}CO.

Altogether, ^{13}C-edited experiments might have been underexploited for in-cell structural biology by NMR. ^{13}C-labeling is most often more expensive than ^{15}N-labeling, but the advantages of 2D ^{1}H-^{13}C correlation spectra show promising results, which could be further improved by some recent isotope labeling methods and pulse sequences.[438,439]

2.3.4. ^{19}F-NMR spectroscopy

Because fluorine is almost absent from the biological realm,[440] in-cell ^{19}F NMR is naturally free from any cellular background signal. For the same reason, it is also not straightforward to produce samples generating ^{19}F-NMR signals of interest. As mentioned earlier, these can be generated by incorporating non-natural amino acids or nucleic bases, either using recombinant production or chemical reactions or synthesis (see chapter 2.2.1.). ^{19}F-NMR has two important advantages: i) ^{19}F has a high gyromagnetic ratio, which ensures high sensitivity *a priori*; ii) the magnetic resonances of ^{19}F nuclei are extremely sensitive to any subtle changes in their local environment, which produces a chemical shift range ~100-fold broader than that of ^{1}H.[117,402,441] Hence, ^{19}F-NMR investigations can be executed by recording simple one-dimensional spectra, which favors sensitivity. (see chapter 2.1.3.). This makes ^{19}F an appealing probe for nailing structural changes or dynamics of macromolecules with multiple stable conformations,[402,442] as shown by a number of recent important reports on GPCRs for example.[443–447] However, ^{19}F-NMR has its Achille's heal: the high chemical shift anisotropy (CSA) of ^{19}F nucleus provokes fast T2 relaxation, hence broad linewidth and reduced signal intensities, a problem that worsens as tumbling time and/or magnetic field increase.[117,402,441] As mentioned earlier, proteins/nucleic acids tumble slower in cells, due to cellular viscosity and multiple interactions. ^{19}F-NMR has had interesting contributions though, and we provide below some technical details about their spectroscopic aspects.

The first in-cell ^{19}F-NMR spectra were recorded by Brindle and colleagues in the late 2000's.[141,142] Four different enzymes were tested, the samples being generated from overexpression in yeast strains incorporating 5-fluorotryptophane (5F-Trp). Two of them (45 and 104 kDa) kept sufficient intracellular mobility to be observable by ^{19}F-NMR, while the remaining two enzymes (a 210 kDa homotetramer and a 98 kDa dimer) were apparently immobilized and not observable by solution NMR. Even for the observable enzymes, one-dimensional in-cell ^{19}F-spectra required 2000 scans to yield sufficient S/N at 400 MHz (^{1}H-NMR frequencies) although they reached 0.4-1 mM intracellular concentrations. This was due to faster T2 relaxation in cells, provoking also a ~4-fold increase in linewidth as compared to *in vitro* spectra. Brindle and colleagues used a "Ernst-

angle" to improve the S/N-PUT; acquisition times were about 20 minutes.

More than ten years later, Li, Pielak and their colleagues reported their first in-cell ^{19}F-NMR investigation,[193] starting a series of studies constituting the main body of knowledge established on this approach. Overall, they carried out their works using: i) ^{19}F-dedicated (cryo)probes at 500-600 MHz; ii) one-dimensional ^{19}F-NMR spectroscopy accumulating ~2000 (room temperature probes) or ~500 (cryoprobes) scans in about one hour or 15 minutes, respectively, using 90º pulses; iii) ^{1}H-optimized (cryo)probes at similar magnetic fields to compare or complement ^{19}F-spectra with 2D ^{1}H-^{15}N HSQC or TROSY spectra using similar acquisition times (see chapter 2.3.2.). Their in-cell samples were either *E. coli* cells incorporating fluoro-Phe/Tyr/Trp and overexpressing a protein of interest, either *X. laevis* oocytes microinjected with a recombinant purified fluoro-Phe/Tyr/Trp-protein. These present intracellular concentrations of ~0.5-1 mM or ~0.1 mM, respectively. The fluoro-amino acid incorporation during recombinant production was not necessarily stoichiometric though.[448]

Importantly, the incorporation of fluoro-amino acids permitted the ^{19}F-detection of small (~10 kDa) proteins like ubiquitin and calmodulin in cells, which did not yield any ^{1}H-^{15}N signals because of the many transient, unspecific interactions with cellular components.[192–195] Li, Pielak and their colleagues showed convincingly that ^{19}F-NMR was capable of detecting a panel of fluoro-proteins previoulsy not observable using the more standard ^{1}H-^{15}N or ^{1}H-^{13}C NMR spectroscopy, both in *E. coli*[189] and *X. laevis* oocytes.[286] Nevertheless, some correlation exists between ^{1}H-^{15}N detectability and ^{19}F-signal linewidth.[189,192–195,197,199,201,286]

Indeed, Li, Pielak and colleagues reported systematically broad ^{19}F-signal linewidths from fluoro-proteins in cells: these are soon > 0.5-1 ppm large, even for <10 kDa proteins when they experience abundant intracellular interactions.[193–195] Consequently, in-cell ^{19}F-NMR peaks overlap very often, even though fluoro-Phe/Tyr/Trp resonance are naturally well-dispersed.[193–195] This depends on the intracellular T2 relaxation, which is linked to tumbling times and interactions: ^{19}F-NMR is very much affected by the interactions with cellular components, like ^{1}H-detected/^{13}C- or ^{15}N-edited NMR (chapters 2.3.2. and 2.3.3.).[195] Sample homogeneity is not an important factor for ^{19}F-linewidth in *E. coli* samples,[195] at the opposite of what was observed in oocytes.[195,286] However, the more crowded *E. coli* cytosol provokes higher viscosity and more abundant interactions than in frog oocytes, so that similar ^{19}F-linewidths are measured for small proteins in these two cell types (Figure 15).[286]

Hence, broad in-cell ^{19}F-linewidths call for a cautious choice between fluoro-Phe, -Tyr, or -Trp: it is notably dictated by their number in the protein of interest, and by the dispersion of the corresponding ^{19}F-peaks, in order to avoid peak overlaps and the acquisition of pointless spectra. The fluoro-amino acids used for in-cell NMR studies were 3-fluoro-tyrosine (3F-Tyr),[189,192–197,286] 3-fluoro-phenylanaline (3F-Phe),[189] 6-fluoro-tryptophane (6F-Trp),[195,196,202] 5-fluoro-tryptophane (5F-Trp)[159,197–199,201,278,286] (Figure 16). Crowley and colleagues made the use of 5F-Trp labeling popular in the last years, notably because it is naturally produced by *E.*

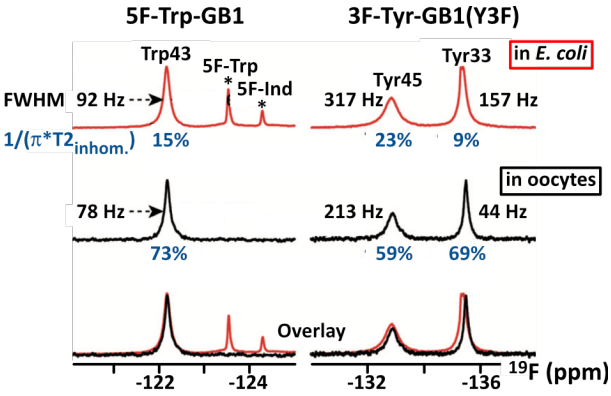


Figure 15: One-dimensional ^{19}F-spectra of 5F-Trp-labeled and 3F-Tyr-labeled GB1 in *E. coli* (red) and in oocytes (black). The Full Width at Half Maximum (FHWM) is indicated for every peak. The contribution of sample inhomogeneity to peak broadening is indicated in blue ($1/(\pi*T2_{inhom.})$, as measured from the difference between the ^{19}F-linewidth and the measured ^{19}F-T2. Stars indicate the fluorinated metabolites remaining in *E. coli* cells (5F-Trp and 5F-Indole). Adapted from ref [286]. Copyright from 2015 John Wiley and Sons.

coli cells from the inexpensive 5F-indole precursor; 5F-Trp is then readily incorporated in the overexpressed, recombinant proteins.[116]

Among the unnatural fluoro-amino acids used for in-cell NMR, the trifluoromethyl-phenylalanine (tfm-Phe) can generate the sharpest signals because of the higher mobility and reduced CSA of the trifluoromethyl function.[193,449] To the best of our knowledge, it was not used anymore for in-cell NMR after 2010, probably because its insertion using amber-codon recoding is more tedious and provides lower yields. It had nevertheless the great advantage of incorporating a ^{19}F probe at a single, well-defined position, and allowed the detection of large proteins in *E. coli* (up to 100 kDa), which were otherwise invisible using ^{1}H-^{15}N NMR spectroscopy.[193]

In-cell ^{19}F-NMR looks rather useless for IDPs: the chemical shift dispersion is poor, and the peak broadening in cells is such that they readily overlap, as shown with the model IDP α-synuclein.[193,286]

^{19}F-NMR provided promising results for studying nucleic acids in cells recently. Xu and colleagues introduced a 3,5-bis(trifluoromethyl)benzene moiety at the 5' termini of DNA oligonucleotides, which were later delivered in *X. laevis* oocytes or in HeLa cells.[134,292,329,450] They yielded broad, but detectable signals in 1 hour for intracellular concentrations of about 150 µM, at 400 MHz using a room-temperature probe. Srivatan and colleagues incorporated a 5-fluorobenzofuran-2'-deoxyuridine in their oligonucleotide, which is much less mobile, and thus suffers of very fast T2 relaxation in cells: it required 8-10 hours of acquisition times to detect very broad ^{19}F-signals of their oligonucleotides at ~180 µM in *X. laevis* oocytes, using a 600 MHz spectrometer equipped with a room temperature probe.[293]

Potential issues are often hypothesized about the impact of unnatural fluoro-amino acids or -nucleic acids on proteins/nucleic acids conformation and stability. On average, they tend to faintly stabilize protein folds, but this depends naturally on the type of fluoro-amino acid.[451] As far as in-cell NMR is concerned, only aromatic fluoro-amino acids have been used. These have been shown to be capable to estab-

lished new hydrogen bonds through the fluorine atom, and to affect the structure in some cases.[402,441,451–453] Consequences on enzymatic capacities have been reported, notably because fluorination affects tyrosine pKa.[451] Some other proteins are not perturbed by the incorporation of fluorinated aromatic amino acids.[451,454–456] The accumulation of numerous fluorinated amino acids in hydrophobic cores appears to have destabilizing consequences, and it is advisable to dilute fluorine incorporation, notably for the more abundant phenylalanines.[402,441,451,453] Most of mono-fluorinated nucleotides are thought to have only minor consequences on nucleic acids thermodynamic stability, whereas tri-fluorination might be more deleterious.[135,136,138,457]

Altogether, in-cell ^{19}F-NMR permitted the detection of proteins, whose ^{1}H-^{13}C/^{15}N signals were broadened beyond detection. It has been however limited to the study of rather small proteins, because of the rapid intensity losses suffered at higher molecular weights. Hence, we have seen that the large ^{19}F CSA provokes fast T2 relaxation and very broad signals of larger proteins and nucleic acids in cells (Figure 16). All these observations were reported from studies using conventional one-dimensional ^{19}F-NMR pulse sequences. Recent reports have shown that this large CSA can be used advantageously if ^{19}F-^{13}C spin pairs are created, both in aromatic amino acids and nucleotides:[118,119,137,139,410] these pairs permit the selection of so-called TROSY components of the ^{19}F/^{13}C resonances, which have slow relaxation and narrow linewidths even for high molecular weight species. This approach might have interesting applications for in-cell NMR studies in our opinion, especially for nucleic acids.

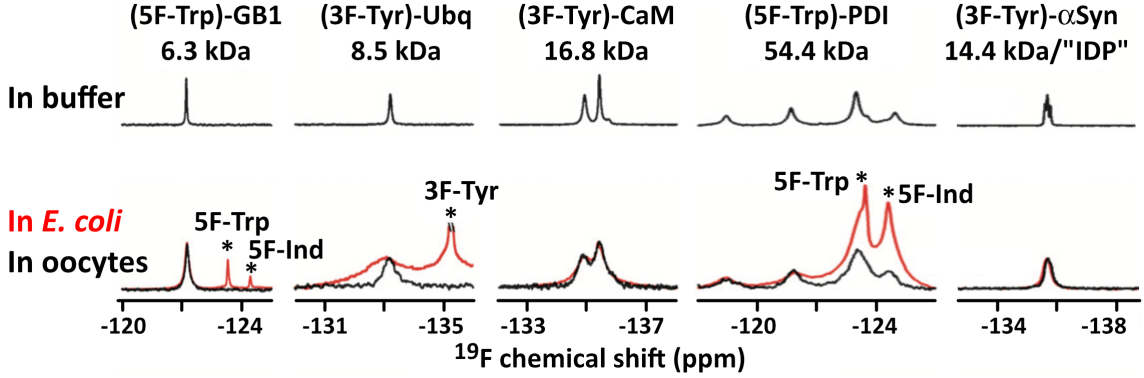


Figure 16: ^{19}F NMR spectra of recombinant proteins incorporating either 3F-Tyr or 5F-Trp. Spectra recorded *in vitro* show the progressive broadening accompanying the increase in molecular weight and tumbling time. Proteins were injected in oocytes at final concentrations of 100-200 µM; intracellular concentrations in *E. coli* are less precisely characterized, in the millimolar range. Spectra were acquired at 600 MHz on a ^{19}F-optimized cryoprobe in ~45 minutes for GB1 and α-synuclein (αSyn), 1.5 hour for ubiquitin (Ubq), 3 hours for calmodulin (CaM) and protein disulfide isomerase (PDI). Spectra recorded in *X. laevis* oocytes and *E. coli* are in black and red, respectively. In *E. coli* spectra, peaks of free metabolites 5F-tryptophane (5F-Trp) and its precursor 5F-indole (5F-Ind) yield sharp peaks highlighted by stars. Adapted from ref [286]. Copyright 2015 John Wiley and Sons.

2.3.5. Distance measurement and structure determination

Two types of distances have been measured using in-cell NMR spectroscopy: i) the internuclei distances via the nuclear Overhauser effects (NOEs); ii) the distance between a nucleus and a paramagnetic center via the pseudo-contact shifts (PCS).[458] A third type of information for structural calculation comes from backbone chemical shifts and their correspondence with backbone torsion angles. Ito, Güntert and colleagues included these angles as restraints in structural calculation, as predicted from chemical shifts by TALOS-N.[168,169,227,310,459] Waudby and colleagues have also used backbone chemical shifts to quantify the secondary structure propensities of the model IDP α-synuclein in *E. coli*.[175] We will not develop further this aspect, which is not specific to in-cell NMR.

NOEs reveal their close distance in space between nuclei (~5-6 Å in most cases). They are traditionally detected by crosspeaks in 2D or 3D NMR spectra, whose intensities are correlated to $1/d^6$, with d the distance between two nuclei. In-cell NOEs have been measured and exploited by Ito, Güntert and their colleagues. We presented the experimental details of their 3D ^{1}H/^{13}C/^{15}N spectra acquisitions in 2.3.3. The analysis and exploitation of the recorded datasets requires also some adaptations, which benefited from the most advanced techniques for i) combined automated NOEs assignment and structure calculation and ii) structure refinement, as developed by Güntert, Ikeya and their colleagues.[460,461] Indeed, 3D spectra in a cellular background produces i) broader peaks than their *in vitro* counterparts, and ii) peaks from cellular components that should not be integrated in the later structure calculation. These aspects had already been recognized in 2006 by Wüthrich and colleagues, who attempted to determine a structure in cell extracts.[462] Broad peak linewidths are deleterious for NOE-based structure determination.[463,464] Overlapping peaks are either useless or even misleading; broad peaks generate important assignment difficulties, ambiguities and possibly erroneous distance constraints. Moreover, the 3D spectra used for side-chain resonance assignment (HBHA(CBCACO)NH, H(CCCO)NH, and (H)CC(CO)NH) are affected by the faster T2 relaxation in cells. Consequently, a number of expected magnetization transfers and their resulting peaks could not be detected in the 5-6 hours allocated before cell leakage, and the corresponding assignments could not be achieved.[169,226] Ito and colleagues proceeded in two different ways: i) from 2016, they used an automatic peak assignment procedure in the NOESY spectra, where the expected intra-residue crosspeaks appear;[169] ii) in 2019, they prolonged cell viability by using a flow-probe bioreactor, which permitted longer acquisition times and improved detection of crosspeaks in 3D spectra.[227] Moreover, they

produced a number of samples with various amino acid selective ^{13}C-labeling (AILVTFWY) and assigned manually side chain resonances from ^{15}N- and ^{13}C- separated NOESY spectra (they used this second strategy also in 2009-10, but on deuterated proteins incorporating ^{1}H-^{13}C-Ala/Val/Leu).[167,168,227] Importantly, the authors deplored a considerable number of nOe crosspeaks from the cellular background. This forced them to treat separately the isolated and overlapping peaks, the intensities of the latter class being measured only after background subtraction.[227] Finally, the last steps of structure determination by NMR include refinement calculation driven by physical force fields: in the case of in-cell samples, the sparse and ambiguous data produce models that require algorithms with larger radius of convergence and conformational search than those used for *in vitro* data. Ikeya, Güntert and their colleagues implemented a Bayesian inference-assisted calculation integrating experimental NMR data and torsion angle, replica exchange molecular dynamics. Altogether, these methods permitted to extract as much information as currently possible from the intrinsically mediocre NMR distance measurements in cells. It yielded well-converged structure ensembles (average backbone RMSD ~0.5-0.8 Å for three 7 kDa proteins) (Figure 17). In-cell NOE-derived structures can thus be derived for 10 kDa proteins, and probably up to 15 kDa at least.

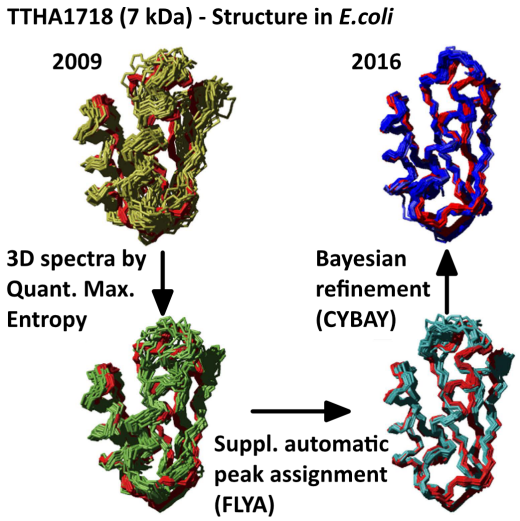


Figure 17: Successive structural ensembles obtained by integrating advanced data treatment, from 3D spectra reconstruction of sparsely sampled datasets by Quantitative Maximum Entropy algorithms, to automatic assignment of side chain resonances by the FLYA algorithm, and to inferential structure determination by the CIBAY module of CYANA. The depicted structures were obtained without the hydrogen bond restraints, which had been used initially in 2009: the authors argued that these restraints impose secondary structure hydrogen bonds from *in vitro* standards, which might not hold true in cells. Adapted from ref [169]. Copyright 2016 from Ikeya et al. under the terms of a Creative Commons CC BY license.

2.3.6. Paramagnetic centers for distance and orientation measurements, and structure determination

Pseudo-contact shifts (PCSs) are present in molecules chelating a paramagnetic moiety, whose unpaired electrons distort the local magnetic field in an anisotropic fashion (typically lanthanide ions). This provokes chemical shift perturbations, the PCSs, of the surrounding nuclei of the same molecule. Remarkably, these perturbations depend on both distance ($1/r^3$) and orientation between the paramagnetic center and the observed nuclei. Dedicated publications explain theoretical and practical aspects of this phenomenon.[465–471] PCSs are thus a rich source of structural information, with two highly attractive strengths: i) they can provide long-distance information, commonly up to 5-6 nm, provided that the chelated center mobility is constrained enough in the studied molecules (recent, improved lanthanide cages might give access to distance measurements up to 20 nm!);[472–475] ii) they can be measured from 1D or 2D spectra, whose acquisition are much faster and more sensitive than those of the 3D spectra necessary for determining NOEs (Figure 18A-B). This makes PCSs very appealing for in-cell NMR studies. However, most proteins are not natively chelating lanthanides. A regular strategy consists in attaching a lanthanide-cages to the macromolecule of interest. Establishing disulfide bonds between a single cysteine and a lanthanide-chelating moiety is a popular method, but it is not appropriate the reducing intracellular environment.[474,476] In 2016, three independent groups sought to use nonreducible thioether bonds to link ^{15}N-labeled, small model proteins (GB1 and ubiquitin) and their home-developed lanthanide-chelating moieties, which made it possible to measure PCSs in live cells (Fig 18).[287,288,362] Two of them delivered the resulting constructs in *X. laevis* oocytes, the third one in HeLa mammalian cells, at final intracellular concentrations of 50 µM. PCSs were readily measured in ^{1}H-^{15}N HSQC (oocytes, at 600 MHz) or SOFAST-HMQC (HeLa cells, at 900 MHz) spectra in 2-5 hours and 30 minutes, respectively. PCSs were highly similar to those measured *in vitro*, which was expected from proteins that have very stable folds. These PCSs were even used for structure calculation, using the program GPS-Rosetta, which integrates backbone chemical shifts and PCSs as structural constraints to generate models.[467,477] Using PCS from 9 samples (3 cysteine mutants, and 3 lanthanides, namely Tm and Tb, plus Lu for the diamagnetic reference, chelated by the DOTA-MPy7 tag), Theillet, Häussinger and colleagues found out that the 100 lowest energy models (among 10,000) of GB1 in cells had an average Cα RMSD of 1.04 and 1.85 Å when incorporating or not PCSs constraints, respectively.[287] Su, Huber, Li and colleagues came to similar conclusions using 8 samples (2 cysteine mutants and 4 lanthanides).[288] GB1 is of course a small model protein, whose *ab initio* structural prediction was already satisfying (in times where deep-learning methods were not even existing…).

To the best of our knowledge, no other in-cell NMR studies were carried out using PCSs, except a short report of PCSs provoked by Tb^{3+} complexation of ^{19}F-labeled calmodulin in *X. laevis* oocytes.[192] The ligation of these first thioether lanthanide tags required rather harsh conditions, i.e. 24 hours long incubation times at 35-40 ºC or one week at room temperature. This was feasible with GB1 or ubiquitin but clearly problematic for most proteins of interest. Recently, Häussinger and colleagues designed advanced tags, which react on cysteines at room temperature and in 30 minutes.[475,478] They combine all the expected qualities: i) their DOTA scaffold ensures good stability and no leakage in cells; ii) their scaffold is conformationally rigid but not too hydrophobic; iii) their linker is also very rigid. Hence, they can create large anisotropic tensors, i.e. large PCSs at long distances. These will certainly have applications in the fu-

ture. It is noteworthy to mention that PCSs have proven to be also useful with ^{13}C-methyl labeling schemes,[479,480] ^{13}C-labeled backbone carbons,[481] or ^{19}F-amino acids, and that recent softwares have been released for improved structure determination using PCS restraints.[479,482,483]

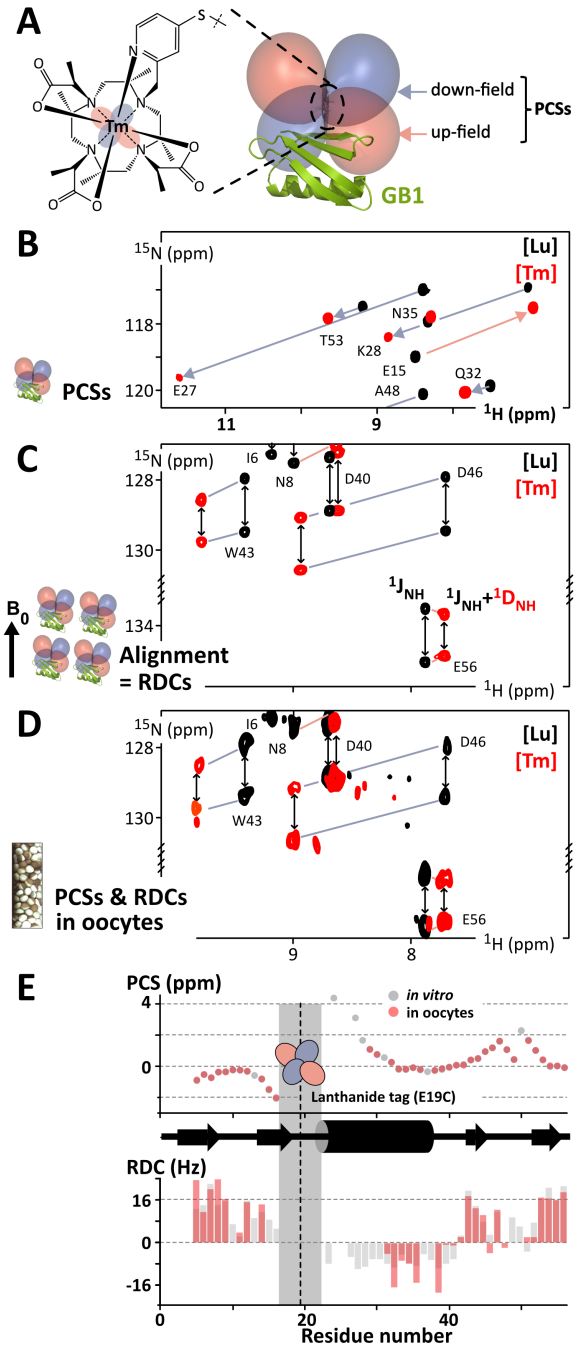


Figure 18: Operating principles of PCSs and RDCs measurement. **A)** Representation of a lanthanide-chelator DOTA-M7Py, which is attached to a cysteine side chain of the studied protein; this figure shows the example of the single point cysteine mutant GB1-E19C- DOTA-M7Py[Lu/Tm]; Lu(III) serves as the reference diamagnetic lanthanide; Tm(III) is paramagnetic with a strong anisotropic tensor; **B)** PCSs are extracted from the difference of chemical shifts observed in the dia- and paramagnetic situations, using classical ^{1}H-^{15}N SOFAST-HMQC or HSQC experiments; **C)** RDCs are extracted from the difference of scalar coupling measured in the dia- and paramagnetic situations, using ^{1}H-^{15}N IPAP-HSQC spectra; **D)** PCSs and RDCs can be measured in cells, here in microinjected *X. laevis* oocytes, using the same ^{15}N-filtered experiments; **E)** PCSs and RDCs were very similar *in vitro* and in cells for GB1. Adapted from ref [287]. Copyright 2016 American Chem Society.

Moreover, rigid lanthanide-chelators can provoke a partial alignment of the protein they are linked to, which permits the measurement of residual dipolar couplings (RDCs) (Figure 18C-D).[474,484] These are also a source of structural information: they report on the orientation and conformational dynamics of ^{1}H-^{15}N amide bonds, with no distance limits.[485] The DOTA-MPy7 tag was actually good enough to provoke a sufficient alignment of GB1 and give rise to RDCs in *X. laevis* oocytes, measured using a RDC-dedicated version of ^{1}H-^{15}N HSQC.[486] These were again similar to those measured *in vitro*, reaching impressive values ranging from -15 to +25 Hz at 600 MHz. As structural constraints, RDCs are less potent than PCSs, so that their addition did not have a strong impact on PCSs-assisted structural calculation.[287] RDCs due to paramagnetism scales with the square of the magnetic field.[466,469,473] RDCs might thus be an important source of information on structure and dynamics of proteins in cells when measured at higher fields, say 900 MHz for example.

Paramagnetic centers can also be used to measure intra- and inter-molecular distances via the paramagnetic relaxation enhancements (PREs). These are generated by a paramagnetic moiety attached to the protein/nucleic acid of interest: unpaired electrons provoke a $<1/d^6>$ distance-dependent dipolar relaxation with the observed nuclei in the ~1.5-3 nm range.[466,487] These can be measured from 2D spectra, and represent interesting restraints for structure calculation or to detect low-populated conformations.[487–489] To the best of our knowledge, PREs have been used only once to characterize in-cell distances on the model IDP αSyn in mammalian cells by Theillet and colleagues.[176] At the opposite of PCSs-tags, PRE-dedicated paramagnetic species have preferably isotropic magnetic susceptibility tensors, like Gd(III), Mn(II) or nitroxide functions. These must be attached at a precise localization on the protein to obtain PRE-derived average distances between the paramagnetic center and every observed nuclei. The typical strategy for *in vitro* studies is to create disulfide bonds between a cysteine side chain and a lanthanide-chelator or a chemical function carrying a nitroxide radical. In-cell studies necessitate non-reducible linkage, like thioether bonds, and the most common nitroxides have limited intracellular lifetimes, even though improved probes have been designed.[490,491] Theillet and colleagues have attached a DOTA-derived cage chelating Gd(III) (paramagnetic) or Lu(III) (diamagnetic) on cysteine side chains via a thioether bond using a maleimide conjugation. The product proved to be extremely stable in cells, with no leakage observed after 24 hours.[176] Peak intensities were measured in 2D ^{1}H-^{15}N L-HSQC recorded on αSyn carrying para- or diamagnetic tags at ~100 mM ($t_{acqusition}$~7-8 hours); the peak intensity ratios gave access to residue specific PRE values, using the treatment proposed by Wagner and colleagues.[492] Proteins carrying both dia- and para-magnetic tags were simultaneously delivered in cells, which permitted to derive internal intensity references: the 6-10 residues neighboring the cysteine-DOTA[Gd(III)] are broadened beyond detection, so that the detected peaks of these residues correspond to the diamagnetic counterpart and help to normalize the delivered intracellular concentrations and NMR intensities.

The following calculation of average PRE-derived distances requires several input parameters: i) the PREs; ii) the ^{15}N-T2 and $^{1}H_N$-T2 in the diamagnetic form; iii) the correlation time of the couple {unpaired electron – observed $^{1}H_N$}. The latter values were approximated by assuming that i) DOTA[Gd(III)] and the residue-specific had non-correlated intramolecular motions (αSyn is an IDP); ii) the intracellular viscosity was linearly slowing down the reference correlation time of the DOTA[Gd(III)] cage; iii) the intracellular $^{1}H_N$ correlation time was calculated from the measured ^{15}N T1/T2/het-nOes (see chapter 2.3.8.). The interpretation of the PRE-derived average distance is still controversial: i) it relies on many experimental parameters; ii) it uses on a simplified model adapted to simplified Solomon-Bloembergen equations; iii) most importantly, PREs report distances according to a biased average $<1/d^6>$, which overweighs the compact states in the conformational ensemble.[47,493,494] Solving structures using in-cell PREs would clearly require more advanced data analysis.[482,495,496] PREs have a last important application though: they can inform on intermolecular proximities. For example, Theillet and colleagues attached a DOTA[Gd(III)] tag on ^{14}N-αSyn cysteine mutants and observed the variations in peak intensities on the population of WT ^{15}N-labeled αSyn: lower intensities in some regions revealed transient intracellular ^{14}N-αSyn:^{15}N-αSyn interactions.[176]

We must mention the fact that choosing appropriate lanthanide-tags is a trial-and-error process. Similar problems are met when searching proper positions for attaching fluorophores. Cysteines have to be solvent-accessible of course, which does not warrant their equal reactivity though. At some positions, the tags may destabilize the protein fold, or hinder important interactions with cellular binding partners. Moreover, lanthanide-tags can be more or less mobile depending on their position, and consequently, more or less capable to yield measurable PCSs and RDCs. Once more, the preliminary *in vitro* characterization of purified material is a prerequisite.

Altogether, the feasibility and usefulness of paramagnetic probes are clear for in-cell structural studies. They confer the ability to produce information on tridimensional structures using fast and sensitive 2D NMR spectra. They may have applications to derive structural information from proteins/nucleic acids showing multiple stable conformations and to investigate allosteric mechanisms in cells. They may also contribute to intracellular studies of multiple domain proteins or protein:RNA complexes, as shown in studies on purified material *in vitro*.[497–499]

2.3.7. Intracellular mobility

Cellular environments alter the rotational and translational motion of molecules. These two mobilities are similarly affected at the nanometer/nanosecond scales, but translational diffusion is restrained by cellular macromolecules and organelles at the micrometer/millisecond scale.[13] The tumbling time of a given molecule can also be different at the nano and at the milli-second timescales, in function of the (un)specific and transient interactions that it experiences in cells. Rotational and translational correlation times are thus often decorrelated at the millisecond time scale. NMR spectroscopy can inform on both phenomena independently.

NMR spectroscopy has very interesting capacities in measuring translational diffusion coefficients, thanks to Pulsed Field-Gradient (PFG) NMR. We refer the reader to dedicated reviews on PFG NMR theory and its numerous applications.[500–502] It relies on the application of strong magnetic field gradient pulses before and after a chosen diffusion delay: this results in a final NMR signal attenuation, which is correlated to the mean distance covered by the population of the observed molecule during the diffusion delay. PFG NMR permits to measure molecular displacements from 10 nm to 100 µm occurring in a few dozens of milliseconds, i.e. the chosen diffusion delay. ^{13}C- or ^{15}N-filters can be added to PFG-NMR pulse sequences, which was first used to measure ^{15}N-labeled GB1 diffusion in *E. coli* cell lysates.[503] To the best of our knowledge, PFG-NMR has been used for in-cell studies only by Waudby and colleagues.[174,175] Their priority was actually not to measure protein diffusion in cells, but to establish a method to filter out the molecules that translated too fast, i.e. those that had leaked from cells and were not restrained in their diffusion by the cell dimensions anymore. They used pulse sequences named ^{1}H-^{15}N HSQC- or BEST-XSTE,[504] and ^{1}H-^{13}C HMQC-STE (also called ^{1}H-^{13}C methyl-TROSY HMQC-STE 3D DOSY-TROSY)[505]. They 300 ms diffusion delays and PFGs at the maximum accessible strength (11.7 T/m) to keep only the signals from proteins entrapped in the micrometer-scale *E. coli* cells. They also measured the diffusion coefficients of αSyn in *E. coli* cells using diffusion times between 11 and 20 ms. PFG-NMR pulse sequences exist that are adapted to large, slow tumbling ^{15}N-labeled macromolecules (^{1}H-^{15}N TROSY-STE),[506] or to ^{13}C-^{15}N-labeled disordered proteins that experience fast water-amide proton exchange at physiological temperature and pH (^{1}Hα-DOSY−^{1}H/^{15}N-HSQC and ^{1}Hα-DOSY−^{13}C′/^{15}N-HSQC)[507]. These have their own advantages that may be useful for in-cell studies.

Rotational diffusion coefficients (also evoked with the related concepts of rotational correlation times, or tumbling times) have drawn more attention in the in-cell NMR community, because it is linked to T2 relaxation time and thus to signal intensity and linewidth. Proteins/nucleic acids tumbling time occurs at the 5-100 ns time scale, and determines T1 and T2 relaxation of ^{1}H/^{13}C/^{15}N/^{19}F nuclei to a good extent.[85,86,402,508] However, T1 and T2 are also affected by ps-ns internal mobility of the observed nuclei, and by µs-ms conformational exchange or interactions. For the nuclei and the NMR frequencies usually probed, T1 is more affected by the ps-ns phenomena, and T2 by the µs-ms ones. The standard approach uses ^{15}N-amide relaxation measurements to gain access to global rotational correlation time and ps-ns internal motion in macromolecules. In particular, the average T1/T2 ratio permits to calculate the global correlation time in the case of isotropic tumbling in a dilute solution.[509] A third value, the heteronuclear nOe ^{1}H-^{15}N, helps to better deconvolve the phenomena occurring at the ps-ns, ns and µs-ms timescale in a residue-specific fashion.[86,508,510,511] Alternatively, one can interpret the product 1/T1*1/T2: it is constant at a given magnetic field for highly ordered residues in large macromolecules (correlation time above 5-10 ns).[512] A lower 1/T1*1/T2 reveals faster tumbling or increased ps-ns flexibility, and a higher 1/T1*1/T2 reveals µs-ms conformational exchange or interactions. Hence, Pielak, Li and colleagues have shown that the product 1/T1*1/T2 was useful to nail transient interactions in crowded media:[513,514] the average T1/T2 ratio is indeed a poor indicator of the protein correlation time and viscosity in complex environments like cells.

Measuring these ^{15}N relaxation parameters is moreover quite demanding in acquisition time: i) the corresponding 2D ^{1}H-^{15}N HSQC-T1/T2 experiments are not very sensitive, and the 2D ^{1}H-^{15}N HSQC-heteronuclear-nOe is worse; ii) series of such spectra including variable relaxation times (at least ~5) must be recorded to obtain relaxation decay curves. To speed up their acquisition, Ito, Shirakawa and colleagues used in-cell samples, in which only the lysines of the model protein TTHA1718 were ^{15}N-labeled. This permitted to record fast and more sensitive 1D ^{1}H-^{15}N-HSQC-T1/T2 spectra, where lysine amide peaks could be identified individually.[167] The product 1/T1*1/T2 of these lysines revealed that TTHA1718 was experiencing unknown interactions in *E. coli*. More recently, Yao and colleagues acquired series of 2D ^{1}H-^{15}N-HSQC-T1/T2 spectra of the model GB3 proteins in *E. coli*, which yielded the same conclusions from the 1/T1*1/T2 product: transient interactions occur with cellular components.[181,515] Gierasch and colleagues recorded 2D ^{1}H-^{15}N-TROSY spectra of GB1 in *E. coli* to analyze the so-called TROSY and anti-TROSY peaks: the ratio of their ^{15}N linewidths informs about the apparent size of the protein, and is not affected by µs-ms events. They also measured ^{1}H T2 relaxation from their peak linewidth in 2D ^{1}H-^{15}N-HSQC spectra; they concluded that these ^{1}H peaks were broader than what was expected from the molecular size extracted from the TROSY/anti-TROSY peaks. This revealed that supplementary relaxation occurred due to µs-ms transient interactions or to magnetic field inhomogeneities. Danielsson, Oliveberg and colleagues measured one-dimensional ^{1}H-^{15}N-HSQC-T1/T2 spectra of uniformly ^{15}N-labeled folded model proteins in *E. coli* or mammalian cells, and also obtained data consistent with transient interactions.[366,425] They also measured T2 relaxation in pulse sequences where they removed the refocusing CPMG train, which yielded similar relaxation rates. Moreover, they evaluated ^{1}H T2-relaxation of these model proteins by measuring their peak linewidths in 2D ^{1}H-^{15}N-SOFAST-HMQC, which revealed similar T2 increase for both ^{1}H and ^{15}N. These observations suggested that interactions at the µs-ms were not responsible of the faster T2 in cells. This would be consistent with the following idea: the measured correlation times from T1/T2 of folded proteins in cells might reveal an average tumbling affected by multiple interactions at the nanosecond timescale with macromolecular assemblies in cells. This would hold true only for proteins that do not have more specific, high affinity interactions with cognate partners in cells.

More comprehensive information was obtained on disordered proteins, which generate higher NMR signal intensities than folded proteins. IDPs could be investigated using 2D ^{1}H-^{15}N-HSQC-T1/T2/heteronuclear-nOe spectra of uniformly ^{15}N-labeled proteins.[176,426] It appeared that ^{15}N T2 relaxation times were very inhomogeneous through the sequence, which reflected notably the existence of residue-specific transient interactions with unknown cellular components (NB: chaperones were later identified to be important contributors of αSyn intracellular interactions[360]). These experiments were carried out at 283K to limit water:amide proton exchange and thus obtain sufficient IDPs NMR signal intensities. Some contributions from rapid water:amide proton exchange are also likely to bias the experimental T2 values of these IDPs.[516]

Although less common, it is possible to evaluate the rotational correlation time from ^{19}F T1 and T2 relaxation. However, ^{19}F T1 is very sensitive to the internal ps-ns motion and the network of close protons, while ^{19}F T2 is very much affected by transient interactions. This was already observed in the first reports using in-cell ^{19}F-NMR by Brindle and colleagues, who noticed that the T2(linewidth)-derived correlation times of the observed enzymes in yeast were 2-3-fold larger than the T1(inversion-recovery)-derived ones.[141] Li, Pielak and colleagues measured these ^{19}F-T1 and T2 relaxation times in a more precise fashion on GB1 in *E. coli* (and ubiquitin in cell lysates): T1 was measured by inversion recovery and T2 using a CPMG pulse sequence.[195,197,199,278,286] They settled a good framework for understanding and interpretating in-cell ^{19}F relaxation times: i) ^{19}F T1 provides a good measurement of the rotational correlation time (or apparent cellular viscosity), provided that its linearity is verified using glycerol:water mixtures; ii) this condition is met for ^{19}F nuclei that do not experience important internal motion on the 20-100 ns timescale, which calls for testing various amino acid positions in the sequence; iii) by comparing with T1-derived correlation times, the larger ^{19}F T2-derived, apparent correlation times give qualitative information on the extent of cellular, transient interactions established by the characterized macromolecule.[195,197,199,286] Similar to what we have seen with ^{15}N-relaxation measurements, interactions with large cellular macromolecules (> 100 kDa) affect more ^{19}F T1 than T2. Importantly, multiple-fold T2 decrease occurs even if only a few percents of a protein population bind macromolecules transiently.[195]

Measurement of translational and rotational diffusions by in-cell NMR is thus an intrinsically difficult task. The NMR methods permit to characterize them precisely for purified, dilute material. However, the multiple cellular interactions and inhomogeneities in cells forbid to interpret in-cell T1 and T2 relaxation data like their *in vitro* counterparts. Statistical models of the cellular proteomes were proposed by Danielsson and Oliveberg recently, which can help to analyze in-cell NMR relaxation data (see chapter 3.6.).[366] Integrating more experimental data and including those from independent spectroscopies will be necessary to better understand the past observations.

2.3.8. Conformational dynamics

NMR spectroscopy can provide information on molecular structural dynamics from the picosecond- to the second timescale. [47,52,85,86,402,508] This capacity is one of the prominent strengths of biomolecular NMR, and its past and present contributions shaped our mental representation of biomacromolecules. This relies on the interpretation of NMR parameters that we presented in previous subchapters, among which T1, T2 relaxation and heteronuclear nOes, or residual dipolar couplings (RDCs).

Classically, RDCs were measured *in vitro* in alignment media, like suspensions of virus particles stretched hydrogels or liquid crystalline phases.[485,517] This strategy is not adaptable to cellular samples. Another approach consists in attaching the studied protein/nucleic acid to paramagnetic moieties showing anisotropic magnetic susceptibility tensors: these can align weakly the protein population and provoke RDCs. These have been measured only once in cells (see chapter

2.3.5.) and were not used for exploring structural dynamics yet.

As mentioned earlier, measuring relaxation data requires high concentrations and extended acquisition times. In-cell samples are not well adapted to such demanding applications, and only few studies got engaged in this task. The complete set of in-cell 2D ^{1}H-^{15}N-HSQC-T1/T2/heteronuclear-nOe spectra was only measured for the model IDP αSyn at ~100 μM in cultured mammalian cells.[176] T1/T2 series comprised 6 spectra measured at variable relaxation delays, each one lasting about one hour; the heteronuclear nOe spectra was ~12 hours long. This was possible because of the high mobility of αSyn residues, typically observed for IDPs. To be correctly interpreted, the amide:water proton exchange had to be included in the fitting of the T2 relaxation curves (see Baum et al. for a detailed explanation of this aspect[516]). These were extrapolated from the exchange rates measured *in vitro* from samples containing crowding agents, using CLEANEX-PM pulse sequences (see chapter 2.3.9.).[518] Although it is better adapted to folded proteins than IDPs,[47,519–521] this scarce accessible information forced Theillet et al. to use a simplistic Modelfree approach,[522,523] which included a global correlation time, and a residue-specific couple {order parameter + internal correlation time} completed by a R_{ex} accounting for μs-ms interactions. These were later integrated to calculate a residue-specific correlation time, according to the model developed by Poulsen and coleagues.[176,519]

^{15}N-T2 relaxation can be further exploited, using the so-called CPMG relaxation dispersion. This approach relies on the quantification of ^{15}N-T2 relaxation using variable refocusing frequencies, whose analysis can give an access to the quantification of low-populated conformations of interactions occurring in the μs-ms timescale.[49,87,524] This timescale is particularly relevant for biological processes, but such studies are also demanding in protein/nucleic acid concentrations and acquisition times. Indeed, they require to record series of 2D ^{1}H-^{15}N-HSQC-CPMG experiments integrating variable refocusing frequencies. Yao and colleagues managed to carry out such measurements on a GB3 mutant overexpressed in *E. coli*. This small model protein experiences little unspecific interactions and generate high S/N NMR spectra in cells. The authors could measure 12 2D ^{1}H-^{15}N-HSQC-CPMG spectra using a constant relaxation delay of 20 ms in a sequence released by Yang and colleagues.[525] These spectra were recorded in 3.5 hours at 308K (E. coli cells start leaking after 5 hours in these conditions according to the authors), but the authors did not report any quantification of the intracellular protein concentration. They also did not give extensive details on their fitting of the CPMG data: they obtain the ^{15}N chemical shifts, the populations and kinetics of an unfolded conformation of GB3 using an in-house program. ^{1}H-^{13}C methyl-TROSY-CPMG and ^{19}F-CPMG experiments exist, which may be useful for future in-cell studies.[442,526]

Overall, conformational dynamics have not been extensively investigated in cells. This is mostly due to the fact that such characterizations require high concentrations and/or long acquisition times that do not correspond to typical in-cell NMR samples. Amino acid-specific isotope-labeling shall be used in the future, which would allow to record one-dimensional relaxation spectra reporting for a limited set of residues/nucleotides. In a number of cases, quantifying a single conformational exchange from the full set of amino acids is not necessary, and reliable information would be obtained from a few residues as well.

2.3.9. Hydrogen exchange

The water:amide proton exchange rates give indications about the local flexibility and solvent accessibility of proteins/nucleic acids, or about the stability of those adopting folded states.[527,528] NMR had valuable contributions in the field,[529–531] and its capacity to derive residue specific information provided good references for the today more popular H/D exchange mass-spectrometry (HDX-MS) approaches.[527,532] Using ^{15}N-filtered pulse sequences, NMR spectroscopy permits to quantify hydrogen exchange in cells. We detail the strategies adopted in the following.

Shirakawa and colleagues were the first to measure hydrogen exchange in cells in 2009: they produced and purified amide-deuterated ubiquitin, which they delivered in cultured mammalian cells using pore-forming toxins and subsequent resealing.[310] After varying incubation times, they disrupted cells in ammonium acetate at pH=5 to stop the water:amide proton exchange,[533] and recorded 2D ^{1}H-^{15}N SOFAST-HMQC experiments in cell lysates (20 hours long acquisition times for a ^{15}N-ubiquitin at ~20-30 μM in 10^{7} cells). They derived the H-exchange rates from the time-resolved build-up curves of residue specific peak intensities in the 2D spectra. This approach may have one default: the delivery method can drive at least a fraction of the protein population in very peculiar cell compartments like endosomes/lysosomes, where protein unfolding and fast H-exchange are likely to occur.

Pielak and colleagues adopted an inverse method:[159] after the overexpression of a ^{15}N-labeled GB1 protein in *E. coli* using a standard H_2O-culture medium, they plunged cells in a D_2O buffer. They quenched the deuterium:proton exchange at specified times in sodium citrate at pH=3, and measured the NMR peak intensities from ^{1}H-^{15}N HSQC spectra recorded in cleared lysates. The high concentration of GB1 permitted rapid acquisitions in 20 minutes. The method is limited to the study of fully soluble proteins. Moreover, the measured H-exchange may be biased by destabilizing effects of D_2O on native *E. coli* proteins, which could themselves affect the stability of the protein of interest. Finally, the intracellular pH has to be accurately determined, because it is a driving factor of H-exchange rates. Pielak and colleagues had to publish corrections because of that matter,[161] which they controlled carefully in the next studies employing the same H/D-exchange method.[160,163,164]

These time-resolved, quenched lysates methods permit to monitor H-exchange in the timeframe of a few hours. Water:amide proton exchange in disordered proteins occurs at much faster rates than folded proteins, in the order of magnitude of 1-1000 Hz in physiological conditions.[529–531] Pielak used a dedicated pulse sequence, named ^{1}H-^{15}N H/D SOLEXSY,[534] to measure H-exchange rates on ^{15}N-labeled αSynuclein overexpressed in *E. coli* cells.[162] This pulse sequence requires the use of a buffer containing 50% H_2O and 50% D_2O, and is usually performed at temperatures below 25 ºC to yield sufficient S/N. It generates 2D ^{1}H-^{15}N correlation spectra and residue-specific information. Other pulse sequences can be executed in pure H_2O (supplemented with ~3% D_2O for the lock) like the ^{1}H-^{15}N CLEANEX-PM

(the most popular to date),[166,518] or the ^{1}H-^{15}N HETex-SOFAST-HMQC.[535] Other recent pulse sequences might be useful, but their sensitivity might be challenged in the case of in-cell samples.[531,536]

Refined quantification of H-exchange should probably avoid some experimental biases: protein/nucleic acid delivery in cells is likely to drive populations of the studied molecule in peculiar organelles, where unfolding occurs (endosomes/lysosomes); experimental procedures using deuterated buffers, unphysiological pH or temperature might affect cell homeostasis and the stability of numerous endogenous proteins, which, in turn, can (de)stabilize the protein/nucleic acid of interest.

We shall give one careful note to end this subchapter: H-exchange are often translated into "protection factors", which are convenient numbers to use.[527,528,531,532] However, these should be corrected to take into account the local electrostatics, as recently shown by Dass and Mulder:[531] at the physiological pH~7.5 water:amide proton exchange are mainly catalyzed by OH^- ions, which are repulsed by acidic patches. Moreover, we can speculate that cellular metabolites or local cellular environments may play the role of attractive or repulsive factors of OH-, which would affect positively or negatively the measured H-exchange rates on the protein/nucleic acid of interest. This might bias conclusions about fold stabilities and solvent accessibility, for example. The interpretation of H-exchange rates in cells should thus be cautious in many instances.

2.3.10. Practical aspects about folded/unfolded equilibria and protein:protein interactions

We want to briefly stress here an important technical aspect: NMR studies of conformational changes and interactions have to take into account the possible variations signal intensity per molecule. To properly quantify protein populations based on their respective NMR peak intensities, it is important to also quantify the relaxation properties of the exploited peaks, or at least their intrinsic intensities at 100% of population A or B. An impeccable example has been provided by Danielsson and Oliveberg in their study of the folded/unfolded populations of $SOD1^{barrel\text{-}I35A}$.[358] Great variations of amide $^{1}H_N$-NMR signal intensities are also observed for disordered regions of proteins upon pH or temperature changes, due to water:amide proton exchange.[428,436,529–531] It is thus advisable to carry out a careful characterization of the intracellular pH and of the signal intensity per molecule for every residue.

2.3.11. Ligand-oriented observation

In-cell NMR studies are also likely to report for protein/nucleic acid ligand-binding by detecting the ligand. There, the isotope-filter is not as important as it was for macromolecules' detection. One-dimensional ^{1}H- or ^{19}F-NMR spectra are usually sufficient for ligand-oriented NMR studies. We refer the reader to dedicated reviews.[401,537–539] Small molecules have indeed much sharper and more intense NMR signals than macromolecules. Hence, they offer a good contrast as compared to the main, broad cellular background signal. Moreover, cellular metabolites and ligands NMR signals are distinguished by their chemical shifts. Two types of approaches have been carried out: i) the time-resolved monitoring of NMR signals from reporters of an enzymatic activity; ii) the characterization of contacts between ligands and protein targets using STD-NMR and TRNOESY, which are best adapted to membrane proteins, and can thus be classified as "on-cell NMR studies".

The first approach is achieved by recording simple one-dimensional ^{1}H- or ^{19}F-NMR spectra. Breeze, Hu and coworkers acquired time series of 3 minutes long ^{1}H-NMR spectra (with double PFG spin-echo for water suppression)[540] to quantify the degradation of meropenem (383 g/mol) at 100 µM mixed with 500 µL of *E. coli* cells at OD=2.5 in a 5 mm diameter NMR tube (Figure 19).[541] This method was later implemented to report for meropenem degradation by *E. coli* cells in presence of various ligands screened in a 96-well plate; the reaction advancement was measured after quenching.[542]

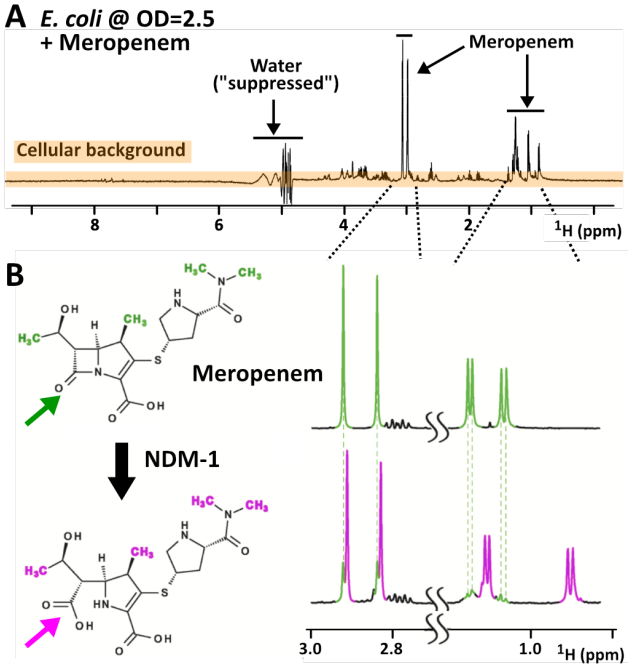


Figure 19: A) One-dimensional ^{1}H-NMR spectrum of live *E. coli* cells at an optical density of 2.5, supplemented with meropenem at 100 µM (water suppression was not ideal, this topic is discussed at the end of this chapter); **B)** ***Left:*** Chemical structures of meropenem before and after hydrolysis by the β-lactamase NDM-1; ***Right:*** close-up views of the ^{1}H-NMR spectra of *E. coli* cells expressing NDM-1 and supplemented with meropenem at an early and a late time point; meropenem and hydrolyzed-meropenem ^{1}H-NMR peaks are highlighted in green and magenta, respectively. Adapted from ref [541] and ref [542], copyright 2014 and 2015 from Ma et al. under the terms of a Creative Commons CC BY license https://creativecommons.org/licenses/by/4.0/.

If the reporters can freely cross the cell barrier back and forth, recording spectra from the supernatant is possible. Hence, Dalvit and colleagues have monitored the degradation of a fluorinated substrate (ARN1203) of the fatty acid amide hydrolase (FAAH) by recording ^{19}F-NMR spectra of the culture medium of mammalian HEK293 cells.[543] After quenching the reactions at different time points, they could store their frozen samples, and later thaw them and record one dimensional ^{19}F-NMR experiments. Acquisition time sof 25 minutes were enough to yield high S/N spectra of ARN1203 at 60 µM. Note that in the case of ^{19}F-NMR, the high CSA of ^{19}F can provoke very different linewidths and

signal intensity per mole for ^{19}F-substrates and ^{19}F-products, when these have distinct molecular weights (Figure 20).

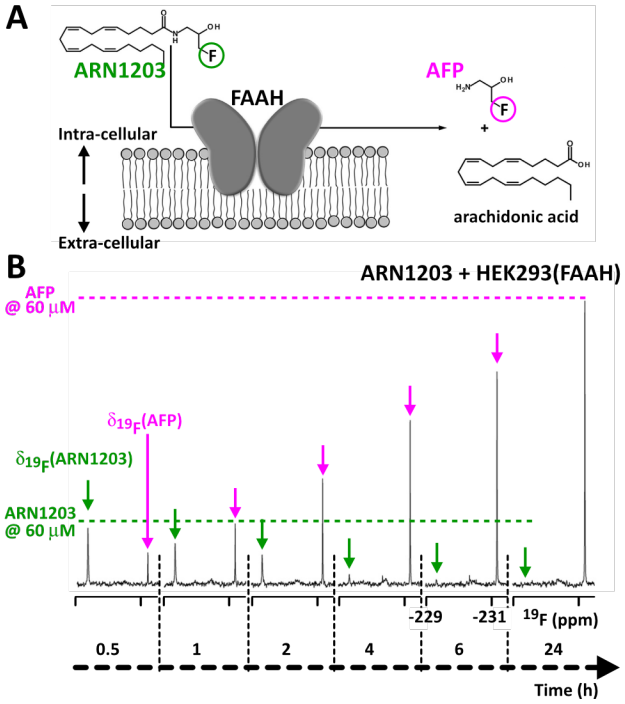


Figure 20: A) Schematic representation of the cleavage of ARN1203 by the enzyme FAAH in the membrane of mammalian cells; **B)** Close-up views of the time-dependent one-dimensional ^{19}F-NMR spectra recorded from quenched supernatant of HEK293 cells stably transfected to overexpress FAAH; The intrinsic signals of ARN1203 and 1-amino-3-fluoro- propan-2-ol (AFP) at 60 mM are highlighted by green and magenta dotted lines, respectively; green and magenta arrowos indicate the time-dependent intensities of ARN1203 and AFP, respectively. Adapted from ref [543]. Copyright 2016 Elsevier.

The second ligand-oriented approach relies on the use of Saturation Transfer Difference NMR (STD-NMR). This technique became popular about 20 years ago in the field of drug research. We refer the reader to dedicated reviews.[537,544,545] It permits to obtain a footprint of the binding epitope, or more simply to detect weak-affinity interactions, with Kd between 10^{-7} and 10^{-3} M, and k_{off} between 10^2 and 10^6 s^{-1}.[537,546] It has been applied to mammalian cells, even though one recent study focused on bacteria.[547] STD-NMR has been used extensively to study drugs and glycans-binding of viral particles, which we will not detail here (see Theillet et al. Prog NMR Spec 2021). Still, it is worth mentioning that these studies have revealed some important technical aspects of STD-NMR using very large particles or cells. First, the off-resonance saturation, which serves to acquire the reference spectrum, must be applied at -300 or +300 ppm, i.e. far from the frequencies used with purified material in dilute solution. Indeed, off-resonance ^{1}H irradiation at -80 or +80 ppm provokes an observable saturation of viral particles, which is transmitted to the ligand of interest; this does not produce a reference spectrum anymore.[548] Second, the global T1 relaxation times of cells is large, so that long interscan delays (≥ 25 seconds) would be necessary to recover the equilibrium magnetization. This would avoid unwanted saturation effects between the interleaved on- and off-resonance saturation scans.[549] A balance might be found between a lower STD contrast and shorter interscan delays to improve the S/N-PUT. Most in-cell STD studies were carried out using interscan delays of 1 to 3 seconds. Third, STD-NMR experiments are more sensitive when performed in D_2O instead of H_2O media: D_2O decreases the transfer of saturation between the irradiated macromolecules and the bulk solvent, which, in turn, permits an improved transfer to the studied ligand.[549,550] D_2O makes T1 relaxation even longer though, which calls eventually for longer interscan delays. STD experiments are generally 1-2 hours long, and appeared to stand deuterated solvents for that amount of time.[398,547,551–558]

Overall, on-cell STD experiments were performed using common pulse parameters: 40 to 60 Gaussion-shaped saturation pulses of 50 ms were applied most of the time, and the experiments last generally 1 to 2 hours. Longer saturation delays would prohibit any epitope mapping.[559,560] The ligand concentration is also rather standard for this type of experiments, we between 0.1 and 3 mM. T1ρ spin-lock filters were eventually used,[547,551] but these are not sufficient to remove the cellular signal background from the STD spectra when the cell density is too high (see in next paragraph). Cell density is actually a key parameter, because it defines the concentration of target proteins. One of the first on-cell STD studies investigated platelets and their 40.10^3 $\alpha_{IIb}\beta_3$ integrin molecules per cell: about 10^{10} cells/mL were resuspended in the NMR tube to reach a receptor concentration of 0.5 µM.[551] Other cells were used for STD studies, which express natively high levels of target receptors, in the range of $20\text{-}50.10^3$ per cell.[552,561,562] Transfected cells can express the protein target to higher levels, i.e. 10^6 per cell,[554,556] so that STD studies were more often carried out using ~$1\text{-}5.10^7$ cells/mL.[398,552–558,561,563–565] The receptor:ligand ratio was usually in the range of $1{:}10^3\text{-}10^5$. The final S/N of these experiments is as dependent on the binding kinetics as their *in vitro* counterparts: k_{off} should be between 10^2 and 10^6 s^{-1}.[537,544,545] STD-NMR studies on cellular samples were mostly performed using one-dimensional ^{1}H-NMR spectra, but Konrat and colleagues reported recently an interesting application of 2D ^{1}H-^{15}N STD-NMR spectroscopy to nail the binding epitope of the ^{15}N-labeled disordered protein osteopontin.[398]

The standard STD approaches disperse the saturation to the whole cellular environment, and cannot restrain it selectively to a chosen cellular protein. This leads to two potential issues. First, this yields an important background signal from the cellular content, when high cell density is used. This hampers the later analysis severely. As a result, early STD-NMR performed on cells were carried out using a first STD-Difference mode (STDD-1): a first STD experiments was recorded in absence of the ligand, the second one with the ligand (Figure 21A).[551,552] A second type of problem emerged: other unspecific contacts can lead to promiscuous STD build-up, e.g. interactions between hydrophobic ligands and cellular membranes. In this regard, it is recommended to work in serum-free medium: weak interactions often occur between serum proteins and the studied ligands, yielding STD signals that can bias the interpretation.[566] A second STD-Difference mode (STDD-2) was proposed: the first STD experiments is recorded using mock-transfected cells (or expressing an inactive mutant), the second one with cells expressing the target-protein of interest (Figure 21B).[553,554,567]

The first STD spectrum is subtracted from the second one, giving rise to a STDD spectrum reporting the specific effects of the contacts between the ligand and the target-protein. This protocol requires two cell samples with very similar compositions, irrespective of the expression of the target-protein. The spectra subtraction decreases the S/N though. It can be avoided if the mock-sample yields a null STD-spectrum, as reported in some other studies.[398,547,555–558] A number of studies have been carried out using cells that naturally overexpress a membrane receptor.[561,563,568–570] There, the results should be taken carefully, unless STD spectra are also carried out in presence of a known ligand of the target-protein.[564,565]

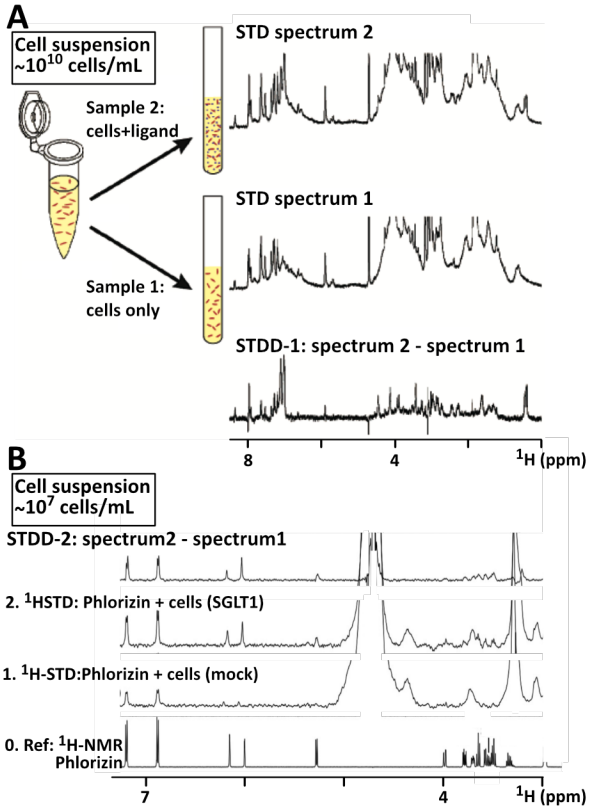


Figure 21: Examples of STD-Difference processing. **A)** In the case of high cell density, the STD-spectrum-1 of cells alone is subtracted to the STD-spectrum-2 of cells in presence of the ligand of interest, adapted from ref [551], copyright 2005 American Chemical Society; **B)** In more normal cases, where about 10^7 cells per mL are used, it is often worthwhile to record a STD-spectrum-1 of the ligand in presence of mock-transfected cells, which is later subtracted to the STD-spectrum-2 of the ligand in presence of cells expressing the target-protein of interest, adapted from ref [553]; Copyright 2011 John Wiley and Sons.

The third technique used to characterize ligand in cellular samples is transferred nuclear Overhauser effect (TRNOE). It relies on the rapid build-up of intramolecular NOEs in small molecules when binding to macromolecules; a rapid exit rate ensures the observation of these NOEs on the free form of the small molecules.[546,571] TRNOEs provide intramolecular distances, provided that short mixing times (<100 ms) are used and that the ligand is in large excess as compared to the protein target.[572] TRNOE has approximately the same requirements than STD-NMR, i.e. affinities in the range of 10^{-6}<Kd<10^{-3} M, k_{off} between 10^2<koff<10^5 s^{-1} , ligand concentration in the millimolar range and target protein in the micromolar range. Hence, TRNOE studies of cellular samples were mostly executed in parallel to STD-NMR studies using the same samples.[652,656,658,660,663–665] They were achieved using NOE mixing times between 100 and 600 ms,[562,573] and the 2D ^{1}H-^{1}H TRNOE spectra were recorded in about 2 to 5 hours.

We shall discuss a few final points. First, all the ligand-oriented ^{1}H-detected techniques yield better results if an excellent and narrow water suppression is executed.[574] It is worth noting that improved water-suppression pulse sequence have been proposed in the last years, which also perform better on inhomogeneous samples.[575–579] As compared to the classical presaturation, DPFGSE and 3-9-19 WATERGATE schemes that are still commonly used by the community, the novel suppression sequences might be fruitfully used on cellular samples. Then, the cellular background should also be suppressed as much as possible. T1ρ filters of a few tens of milliseconds are usually applied, which is sometimes not enough to remove signals from abundant cellular species, including those from lipids. On-cell NMR studies were not very precise about the strength of the applied T1ρ spin-lock, but it is usually $(\gamma/2\pi)B_1$~5 kHz.[580,581] Lower strengths might provoke offset and relaxation-dispersion effects.[582] Interestingly, T1ρ filters could be much longer, up to 500 ms, and thus more efficient without any sample heating, as shown recently.[582] NMR-metabolomics use T2-filters to remove signals from plasma macromolecules for example.[583] The T2-filter have indeed been shown to provoke faster relaxation of macromolecules than the T1ρ filter, and thus a more efficient background removal.[584] However, T2 filters are more difficult to handle, notably because they permit the evolution of J-couplings, even though improved spin-echo schemes exist.[578,585] Longer adiabatic spinlocks can be also achieved up to a few hundred milliseconds, and diffusion filters can be combined to subtract signals from slow-diffusing species like lipids.[586] Longer T2/ T1ρ filters provoke also some relaxation of the NMR signals of interest, so that the best filter delays are certainly case-specific. Finally, cell sedimentation is an embarrassing issue. Most of these ligand-observed studies have been carried out using suspension cells, which settle more slowly than adherent cells. They might be maintained in suspension by spinning tom some extent,[557] or settled on the rotor walls in HR-MAS probes. HR-MAS.[553] Gel-encapsulation is probably the most reproducible approach, as shown recently by Konrat and coworkers.

2.4. The versatility of NMR and its multiplexing capacities

Observing small and large molecules in parallel can be highly beneficial to obtain a complete description of a system, including bulk values like pH or redox potential.[61,160,326,587–589] This is favored by the multi-channel probes, which are nowadays installed on commercial spectrometers. The most common probes for structural studies integrate ^{1}H/^{2}D/^{13}C/^{15}N channels and recent ones can add a supplementary ^{19}F or ^{31}P channel. Hence, it is possible to monitor the same process using multiple complementary readouts, e.g. the levels of substrates, products, co-factors, or cellular metabolites, together with binding events on enzymes or their active/unactive populations. Different spectra can be recorded successively in an interleaved fashion, and

recent systems also permit the simultaneous detection on different channels. It is probably impossible to list all the reports where multiple types of molecules were observed selectively. We selected two prominent examples. The first one is presented in Figure 22: Allain and coworkers have used a $^{1}H/^{2}D/^{13}C/^{15}N/^{31}P$ cryoprobe yielding high detection sensitivity on ^{1}H and ^{31}P channels to monitor the progressive production of small RNAs and their binding to ^{15}N-labeled RNA-binding molecules.[590] This helped to describe in a time-resolved fashion the early liquid-liquid phase separation of these proteins with these small RNAs, and their later re-dissolution at higher RNA:protein ratios. Another interesting example comes from Shimada and colleagues, who recorded spectra of ^{13}C-labeled glutathione and ^{13}C-labeled thioredoxin in parallel to evaluate the intracellular redox potential and the redox equilibrium of thioredoxin (Figure 27).[326]

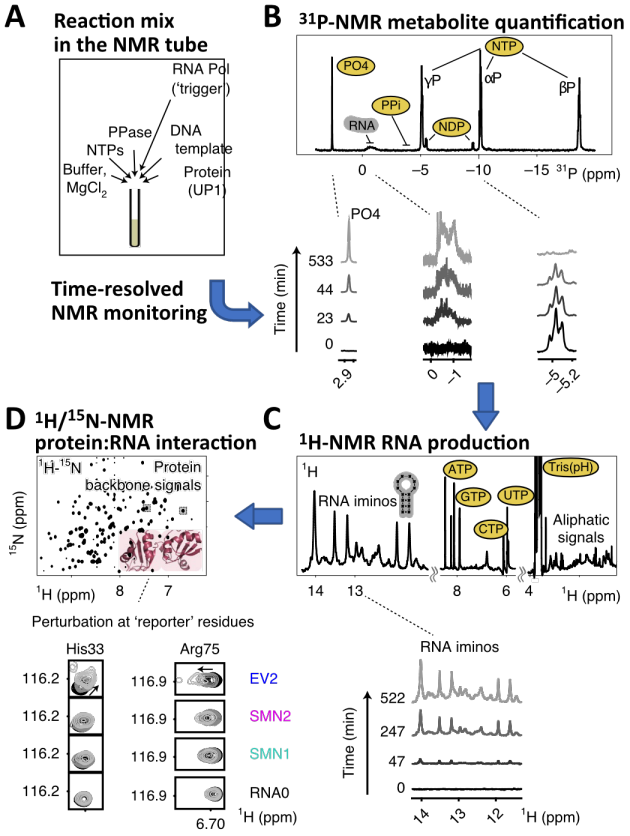


Figure 22: Example of multiple NMR readout. **A)** Here, Allain and coworkers set up a reaction mix in a NMR tube for the production of a small RNA (from a DNA template by the T7 polymerase) and its interaction with a ^{15}N-labeled RNA-binding protein (UP1). The time-resolved NMR monitoring reported in parallel **B)** the NTP consumption via one-dimensional ^{31}P-NMR spectra, **C)** the build-up of RNA production via one-dimensional ^{1}H-NMR spectra, notably by the observation of the imino region (12-16 ppm) and **D)** the interaction of the newly synthesized RNA with the ^{15}N-labeled RNA-binding protein via 2D ^{1}H-^{15}N-NMR spectra, which show chemical shift perturbations due to binding upon the production of the high-affinity EV2 RNA, and much weaker effects in presence of three other RNAs. Adapted from ref [590]. Copyright 2019 Springer Nature.

3. HOW DIFFERENT IS A PROTEIN/NUCLEIC ACID STRUCTURE IN ITS NATIVE ENVIRONMENT? THE NMR CONTRIBUTION.

3.1. Structural stability and dynamics

Debates and theories exist for a long time on the impact of the cellular environment on proteins/nucleic acids structures. Cellular crowding pressure, cellular (un)specific interactions, metabolites or salts can all affect the stability or the structural dynamics of macromolecules.[13,36,37,40,41,591,592] The answers are clearly case-dependent, whether we talk about the type of cells or the individual protein/nucleic acid (not even mentioning cell-cycle phase, stress or subcellular localization). We summarize the information from in-cell NMR studies on that topic below.

We shall make an important warning first. In our opinion, "in-cell thermodynamics" and "in-cell folding or binding free-energy" are misleading concepts. These are forged from theoretical discussions on the impact of crowding or unspecific promiscuous interactions, and hold true only for purified systems testing these hypotheses. Indeed, living cells are not isolated systems and do not let their components evolve towards any thermodynamic equilibrium. Hence, the observed structures and conformational dynamics report for the mixed effects of thermodynamics, kinetics and energy consumption of the multiple actors in a cell. It is thus safer to talk about apparent protein stability, folded/unfolded equilibria, and free:bound ratios. This distinction might sound very conceptual, but it will have practical consequences when in-cell structural knowledge will spur the design of therapeutic applications.

3.1.1. Protein structures determined from in-cell data

Solving protein structures *de novo* in cells may sound odd at first glance: protein folds are not expected to be far different in cells. Folds are encoded in primary structure in robust ways, as recently confirmed by neural networks predicting 3D structures, whose efficiency relies on recognizing structural features through multiple sequence alignments.[4,5] However, how the cellular environment influences local flexibility or subtly reorients secondary structures is unknown. In this regard, structure determination by NMR spectroscopy is an interesting process, in the sense that it results also in the description of local mobilities.

The first tridimensional structure in cells was solved by Ito and colleagues in 2009, using advanced NMR data acquisition and treatment procedures on *E. coli* cells overexpressing a $^{13}C/^{15}N$-labeled small model protein called TTHA1718.[167] They measured and used classical NMR distance restraints, the so-called NOEs, measured between nuclei to calculate structural models. Subsequently, they improved further their data processing and the exploitation of NMR inter-nuclei distances: this helped them to decrease the protein intracellular concentration necessary for in-cell structure determination, down to ~250 µM in *E. coli* cells.[169] Recently, they added information from NMR spectra recorded from supplementary samples including amino acid-specific $^{13}C/^{15}N$-labeling: this clears up the corresponding spectra, allowing the unambiguous assignment of a number of peaks, hence of the corresponding NOEs.[227] This permitted to solve structures of 3 small model proteins (GB1, 7kDa, TTHA1718, 7

kDa, Ubiquitin-3Ala, 8 kDa) at intracellular concentrations of 130 µM in insect cells.

What did we learn from these? First, the folds are conserved, which was the expected results for these three very stable proteins: the backbone RMSD between the mean coordinates of the structure ensembles in cells and in diluted solution are 1.6 Å (Figure 23). However, some interesting differences were observed in the case of GB1 in insect cells: its α-helix appeared to deviate slightly from the solution structure, and one loop show increased mobility. This suggests that the higher mobility of this region is not due to lost internuclei information due to lower sensitivity. Interestingly, this tilt of the a-helix had not been observed in *E. coli*, while the increased loop mobility was present in the structure solved earlier from *E. coli* cells samples.[169] Importantly, the tilt was consistent with local differences in chemical shifts between GB1 in cells and in dilute solution.

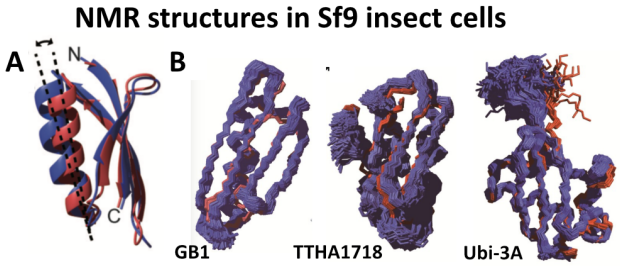


Figure 23: A) Comparison of the lowest energy structure of GB1 in dilute solution (red) and the equivalent in Sf9 insect cells (blue) (the structure in Sf9 cells is strictly speaking the one with highest posterior probability density in the Bayesian inference calculation, the method used to determine structures in cells.). The tilted helix is highlighted by the lines. Adapted from ref [227]; copyright 2019 John Wiley and Sons; **B)** Backbone heavy atoms of the structure ensembles of GB1, TTHA1718 and Ubi-3A determined *in vitro* (purified in dilute solution, 10 lowest energy structures, in red) and in Sf9 cells (20 highest posterior probability density, in blue). Adapted from ref [458]. Copyright 2019 from Tanaka et al. under the terms of a Creative Commons CC BY license https://creativecommons.org/licenses/by/4.0/.

The quality of NMR-derived structures is indeed correlated to the number of structural restraints provided to the model calculation algorithms. The example of TTHA1718 in insect cells might be didactical in this regard: the putative metal-binding loop looks more flexible in cells than in dilute solution, although its chemical shifts are not different. According to Ito and colleagues, this loop may interact loosely with paramagnetic species in cells, which causes intensity losses impeding the detection of NMR peaks and distance restraints in the region.[227,458]

Other groups have attempted to use other structural restraints from proteins tagged with a lanthanide-cages. When covalently attached, these tags can provoke paramagnetic effects that are distance- and orientation-dependent, namely the pseudo-contact-shifts (PCSs), the paramagnetic relation enhancements PREs (PREs), and Residual Dipolar Couplings (RDCs). These can be measured in 2D NMR spectra, which are much faster to acquire and simpler to analyze, and thus give access to a better sensitivity. The proof-of-concept has been shown by three independent groups, which recorded such spectra of ^{15}N-labeled small model proteins (GB1 and Ubi-3Ala) at intracellular concentrations of 50 µM in *X. laevis* and HeLa cells.[287,288,362] The NMR data provided structural information on HN amide functions, and permitted to derive structural ensembles, whose Cα RMSD was 1.05 Å close to the structure determined in dilute solution.

Ito and colleagues managed to record the same NOE spectra from larger proteins (calmodulin, 17 kDa and HRas, 19 kDa) in insect cells, which should enable structure determination. Hence, the method may be valid for ~20 kDa protein domains. The PCSs-RDCs strategy generates lower quality structural models than the NOE-approach. However, PCSs are likely to provide long-distance information, commonly up to 5-6 nm, and RDCs' orientations are distance-independent. Improved lanthanide cages have been designed to allow their attachment at low temperature.[475,478] This should enable future in-cell investigations on mono- and di- or tri-domains proteins, allosteric mechanisms, or on intermolecular interactions.

3.1.2. Folded/unfolded equilibrium

While proteins' folds are not supposed to change drastically in cells, changes in their stability might occur readily. After all, globular proteins are marginally stable objects, whose folding free-energy is commonly about 5 kcal/mol, as measured in their purified, dilute form.[593,594] be more questionable. This corresponds to a handful of hydrogen bonds, which might be altered, compensated or reinforced by cellular crowding or weak interactions with cellular components.[13,36,37,40,41,591,592] These cellular phenomena cannot be addressed by *in vitro* characterizations,[595–597] or computational predictions of protein stability.[598–600] Moreover, these approaches cannot account for the impact of the cellular proteostasis network.[601] At the proteome wide level, poor correlation has been reported between thermal stability *in vitro* and in lysates,[602] and significant differences were observed between thermal stability in lysates and in cells.[603] We are thus in need of experimental data on cellular folded/unfolded equilibria, both for fundamental understanding and for practical use: it has been estimated that about two thirds of monogenic diseases are linked to protein destabilization, for example.[604–607]

NMR spectroscopy provides an immediate readout of structural changes: chemical shifts are exquisitely reporters of local chemical environment, and will thus be markedly different in folded and unfolded conformations. An interesting example was released by Banci and colleagues: ^{15}N-labeled SOD1 disease-mutants overexpressed in cultured human cells generated spectra of unfolded proteins, but the co-overexpression of SOD1-specific chaperone CCS switched the spectra back to those of folded proteins.[255,257] A Similar folding effects were observed upon the addition of a compound, ebselen, which favors SOD1 native dimerization: spectra corresponding to the folded state of SOD1 were obtained in mammalian cells, whose culture medium was supplemented with ebselen.[258]

Danielsson, Oliveberg and colleagues provided more quantitative data on a model β-barrel, destabilized and truncated protein derived from SOD1 (SOD1$^{barrel\text{-}I35A}$). Its folded/unfolded (F/U) ratio was close to unity, and the exchange between the two states was slow (> 1 s). The evolution of this equilibrium was monitored through a temperature array in *E.coli* cells overexpressing ^{15}N-labeled SOD1$^{barrel\text{-}I35A}$, and in culture human cells after ^{15}N-labeled SOD1$^{barrel\text{-}I35A}$ deliv-

ery upon electroporation.[358] The authors exploited the two well-separated NMR peaks generated by folded and unfolded SOD1$^{barrel\text{-}I35A}$(Gln153) in 2D ^{1}H-^{15}N spectra. These took 4 hours to acquire with SOD1$^{barrel\text{-}I35A}$ at ~20-30 µM in about 2-3.10^{7} human cells. These cells were pelleted in NMR tubes without fresh medium circulation, which was probably provoking the observed F/U population drifts at T>293K through the 4 hours experiments. The authors converted the temperature dependent F/U equilibrium in folding free-energy, which permitted to fit folding free enthalpy, free entropy, heat capacity; cold- and hot-unfolding midpoint temperatures were also obtained. They concluded that SOD1$^{barrel\text{-}I35A}$ was destabilized by about 0.5 and 1 kcal/mol in *E.coli* and human cells, respectively (Figure 24). They also established ideal thermodynamic cycles to interpret the variations of heat capacities between *in vitro* and in-cell conditions. The more negative folding heat capacity in cells was related to a larger solvent-accessible area of the unfolded state than *in vitro*; this simplification holds true only in a dilute buffer, while the protein:protein binding heat capacities should not be neglected in cells, in our opinion.[608,609] (NB: They acknowledged these binding contributions recently.[425,427,610]) Moreover, the conclusions of this approach shall be taken with caution: i) the F/U equilibrium of endogenous cellular proteins/nucleic acids is also temperature-dependent and its impact on SOD1$^{barrel\text{-}I35A}$ F/U ratio is unknown;[611] ii) the metabolite content of pelleted cells is not representative of "normal" culture conditions (e.g. while ATP has protein-solubilization properties,[37,592] it is depleted in about 1-2 hours in these experimental conditions[267–269,395]); iii) the ideal thermodynamic cycle can only be closed if a cell is the same with and without the studied protein, and also the same when the protein is folded or unfolded, a model that fits better to ideal crowding agents than to the cellular interplaying proteostasis network. The observed shift of stability in cells is real anyway, and its temperature-dependence was typically following the Gibbs free-energy curves obtained with purified proteins in dilute solution. Interpreting them using the theoretical models from *in vitro* conditions was very tempting, and probably hold some truth.

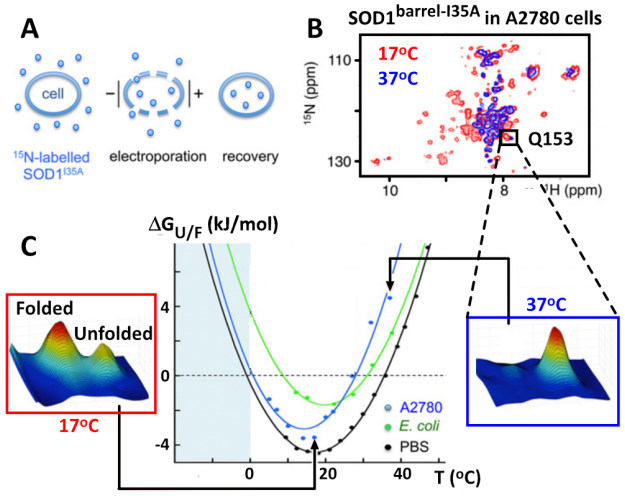


Figure 24: Experimental procedure and analysis proposed by Danielsson et al. PNAS 2015. **A)** ^{15}N-labeled SOD$^{barrel\text{-}I35A}$ (produced from recombinant expression in *E. coli*) was transduced in cultured mammalian cells A2780; **B)** 2D ^{1}H-^{15}N SOFAST-HMQC spectra were recorded at varying temperatures, and gave access to the peak intensities of the residue Q153 in SOD$^{barrel\text{-}I35A}$(folded) and SOD$^{barrel\text{-}I35A}$(unfolded), hence to the intracellular F/U ratio; **C)** These F/U ratios were converted in Gibbs free-energy of folding, and were plotted as a function of temperature, revealing classical $\Delta G_{U/F}(T)$ curves for SOD$^{barrel\text{-}I35A}$ in E. coli, in A278- cells and in dilute solution (PBS: phosphate buffer saline): a temperature of maximal stability exist, due to the negative heat capacity of folding; independently of this representation, *E. coli* and mammalian cells proved to have destabilizing effects on SOD$^{barrel\text{-}I35A}$. Adapted from ref [358]. Copyright 2015 National Academy of Sciences.

As shown above with SOD1$^{barrel\text{-}I35A}$(Gln153), in the case of a two-state model, observing a limited set of resonances can actually provide sufficient information to delineate the folded/unfolded equilibrium. This rationale has been used by Pielak and colleagues to study the stability of two small model protein domains, GB1 and a SH3 domain from *Drosophila* (dSH3), the latter of which has a F/U ratio close to 1 at room temperature. They have used ^{19}F-NMR to observe dSH3 in cells, which integrated a single ^{19}F-Trp residue upon recombinant overexpression in *E. coli* cells supplemented with a fluorinated precursor.[194,198,200] ^{19}F-NMR peaks in folded and unfolded conformations were well-separated. They quantified the two states by integrating the area of these two peaks at varying temperatures,[200] or salt concentrations.[198] They converted these equilibria in unfolding free-energies, which permitted to derive unfolding free-enthalpy, free-entropy and heat capacity of dSH3 in cells. All these values were almost unchanged, as compared to those measured with purified material in dilute solution. [200] At the opposite, exposure to 0.3 M NaCl destabilized dSH3 by 1 kcal/mol in *E. coli* cells, which was reversed by glycine betaine.[198] The authors did not provide any clear numbers on dSH3 intracellular concentrations (probably ~millimolar) and every spectra were recorded in c.a. 20 minutes. They applied recently the same ^{19}F-NMR method to dSH3 microinjected in zebrafish oocytes, where a ~0.5 kcal/mol destabilization was observed.[278]

Although not strictly speaking an in-cell NMR approach, Pielak and colleagues also used a quenched lysate method to quantify water:amide proton exchange on ^{15}N-labeled GB1 destabilized mutants:[158] i) these were overexpressed in *E. coli* cells, which were later pelleted, ii) resuspended in a 100% D_2O buffer, where the H/D-exchange on amides was let to occur through an array of hours-long incubation times, iii) before lysis and H/D-exchange quenching by dropping the pH to 3.5. Finally, 2D ^{1}H-^{15}N NMR spectra of these lysates provided GB1 residue-specific peak intensities, which were inversely proportional to the H/D exchange. The corresponding exchange rates correspond to GB1 conformational breathing and transient unfolding provoking solvent accessibility of the amide function. This method permitted to reveal the cellular destabilization of GB1 in *E. coli*,[158,161] the enhanced effects of some destabilizing surface mutations in cells,[159] the decreased effects of some stabilizing mutations,[164] and its further destabilization at low pH.[163] To what extent plunging cells in D_2O does impact the stability of cellular proteins and their interconnected, promiscuous interactions? D_2O appears to stabilize proteins folding and protein:protein interactions *in vitro*.[612] It is difficult to say whether it affects the GB1 solvent-accessibility in cells. Another question concerns the experimental conditions and cellular proteostasis: when maintained in PBS, without nutri-

ents at 37ºC over hours, *E. coli* cells certainly suffer energy exhaustion, loss of chaperone activity, widespread misfolding, all of which might impact GB1 stability. The set up might be refined with ease.

In 2019 and 2021, Yao and colleagues approached this intracellular conformational breathing in a more quantitative fashion: they measured the so-called "T2 relaxation dispersion" of two destabilized mutants of the small model protein GB3 in *E. coli* cells.[181,515] In the case of a conformational equilibrium between two states at the millisecond time-scale, these experiments permit to quantify i) the populations even if one of them is about 1%, ii) the rates of exchange and iii) the chemical shifts of both populations.[49,87,524] These GB3-derived proteins were overexpressed in the bacteria growing in medium supplemented with ^{15}N-ammmonium salts. This permitted to record the necessary 2D ^{1}H-^{15}N HSQC-CPMG spectra of washed cell pellets in 3.5 hours at 35 ºC (the intracellular GB3 concentration was probably in the millimolar range). The authors could show that the destabilized mutant GB3 had a low-populated conformation similar to that observed in 8 M urea, and that it was about twice more abundant in cells than in dilute solution: 2.6 vs 3.7% for the first mutant, 1.6 vs 2.8% for the second. This was not due to the unfolding rate, but unfolded life-time about twice longer in cells: k_{fold}~2100 vs 1100 s^{-1} for the first mutant, ~2000 vs 1100 s^{-1} for the second.[181] This destabilization was not observed at all in a crowded buffer containing dextran at 100 g/L. At the opposite, it was confirmed in cell lysate but with ~1.8-fold higher k_{unfold} and k_{fold} rates. The same experiments were recorded with cells resuspended in a buffer containing 400 mM of either sorbitol, glycerol, betaine or taurine.[515] The cellular destabilizing effects on GB3 persisted comparison in presence of these osmolytes. However, they affected mostly the intracellular k_{unfold} rates: the presence of glycerol decreased k_{unfold} by 20% (30% *in vitro*), while taurine increased it by 60% (20% *in vitro*); betaine and sorbitol decreased k_{fold} by 15% and 0% in cells, respectively, but increased by 20% and 30% *in vitro*. This poor correlation reveals probably important changes in cell density and macromolecules interactions in cells exposed to these osmolytes. Supplementary transient interactions are reported between GB3 and cellular macromolecules by the average 1/(T1*T2) factor in *E. coli* resuspended in 400 mM betaine or taurine.

Whether these values would hold true for other proteins is unknown. All these results have also to be put in context: they were obtained using cell pellets that were not supplied with fresh medium through the experiments. The resulting ATP-exhaustion holds probably the chaperone machinery in pause, which might actually give access to a chaperone-free quantification of the cellular impact on protein stability. However, as stressed recently by Powers and Gierasch, even only one protein population that misfolds is a matter for the whole proteostasis network, which then loses "bandwidth" to care about the rest of the proteome.[601] Hence, it is desirable to maintain a endogenous proteins well-folded and to introduce as little unfolded populations as possible. It also applies to the unfolded population of the studied protein. This argues notably for performing similar experiments in flow-probe bioreactors, in order to be closer to native conditions in the future.

The relaxation-dispersion experiments are very promising in our opinion. They are limited by two aspects though: they require high protein concentrations in their current version (> 100 µM), and they permit to quantify conformational exchange in the millisecond time scale only. This time scale is actually in the range of native folding events and allosteric regulations. It is interesting to note that ~5 nm large macromolecules explore about 0.5 µm of the intracellular space in one millisecond.[613–616] This represents a large portion of a cell, and that gives probably substantial opportunities to mate with chaperones, membranes, charged or hydrophobic surfaces.

Finally, we shall have a short discussion about how we can interpret the impact of the cellular environment on protein stability. The terms of "quinary structure and/or interactions" are often used. This latter concept emerges from the following credible possibility: natural selection mechanisms would necessarily integrate the effects of the relationship/interactions between a protein and its cellular environment. Crowding, (un)specific interactions with metabolites and macromolecules have consequences on the stability of a protein without any doubts. However, the current in-cell protocols use high concentrations of heterologous, model proteins. These bias the concept of native quinary structure, i.e. a selected protein-specific relationship with its native environment. Although appealing and editorially productive, the concept of quinary structure shall thus be used with caution, in our opinion. Folded/unfolded equilibrium, balanced impacts of crowding and diffuse cellular interactions, those are safer terms, even though less concise. We hope that talented colleagues will find better formulations in the near future.

3.1.3. Nucleic acids

DNA/RNA structure in cells is also at stake in many instances. Among others, RNA are recognized to be very flexible objects,[617–619] and DNA tertiary structures like G-quadruplexes are thought to be of high importance, but known to be highly sensitive to their environment.[620–622] Altogether, the intracellular environments appear to impact nucleic acids structure.[623,624] We will not cover the past analysis of DNA in intact viruses executed by ssNMR (see the following recent reports for an insight in the modern capacities in the field).[625–628] [62] Unless specified, the studies reported below were executed using solution NMR.

In-cell structural NMR studies were first carried out on uniformly ^{13}C/^{15}N-labeled DNA and RNA microinjected in frog oocytes.[289,291,297,629] This approach permitted to acquire 2D ^{1}H-^{13}C and ^{1}H-^{15}N spectra of small oligonucleotides (~15-25 bp) at intracellular 150-250 µM concentrations in 5-8 hours per spectrum. The G-quadruplex structure of $d(TG_4T)_4$,[291] and of $d(AG_3(TTAGGG)_3)$ were confirmed.[629] In the latter case, the nucleic acid concentration was limited to 100 µM for cell survival reasons, which allowed only one-dimensional spectra to be recorded in the imino region. More recently, the global fold of a RNA aptamer was show to be stable after microinjection in frog oocytes at 120 mM using 2D ^{1}H-^{15}N SOFAST-HMQC spectra recorded in about 100 minutes.[294]

Other authors attempted to make profit of ^{19}F-NMR in frog oocytes. Xu and colleagues attached a hexa-fluorinated moiety to the 5' termini of 12-mer RNA, which yielded 3 different ^{19}F-NMR peaks in its 3 conformational states, i.e. monomer, dimer G-quadruplex and two-subunits stacked tetramer G-quadruplex.[292] They recorded one-dimensional

^{19}F-NMR of this 12-mer RNA microinjected at 150-250 µM in frog oocytes in about one hour.[134] Surprisingly, they observed only the ^{19}F-NMR peak corresponding to the tetramer G-quadruplex although it appeared only at millimolar concentrations in dilute solution. Xu and coworkers introduced also a hexafluorinated guanosine in a G-quadruplex DNA oligonucleotide, whose ^{19}F-NMR signal in frog oocytes was at the frequency observed for the G-quadruplex structure *in vitro*.[630] This probe had however G-quadruplex stabilizing properties. This was not the case of another fluorinated probe proposed by Srivatsan and colleagues: they attached a fluorobenzofuran to a 2'-deoxyuridine, which is moreover an environment-sensitive fluorophore.[293] Once integrated in human telomeric DNA oligonucleotides, it generated a signal at different ^{19}F-NMR frequencies in the different G-quadruplex conformations that the oligonucleotides can adopt. This helped to nail the conformations adopted by these oligonucleotides at 180 µM in microinjected frog oocytes in 1D ^{19}F-NMR spectra recorded in about 10 hours. Their one-dimensional ^{1}H-NMR counterparts showed unexploitable, broad signals.

Next, two groups managed to deliver oligonucleotides in cultured mammalian cells in 2018, using either pore-forming toxins or electroporation.[325,364,631] Some leakage have been observed in the case of pore-forming toxins,[325,329,632,633] but this might be fixed by resting and selecting viable cells on plates for a few hours before the NMR acquisition. Alternatively, Katahira and colleagues used a flow-probe bioreactor to remove leaked DNA from the NMR tube.[328] Trantirek and colleagues used electroporation and one-dimensional ^{1}H-NMR in 5 recent publications. The intracellular concentrations of the electroporated nucleotides was about 10-15 µM, which permitted to acquire exploitable 1D ^{1}H-NMR spectra in about 20-40 minutes from pellets of $1.3x10^{8}$ HeLa or HEK cells (they did not keep cells in the NMR tube more than 2 hours).[294,364,365,631] Indeed, this concentration is sufficient to yield enough signal in the imino region of the 1D ^{1}H-NMR spectra (~12-17 ppm). The cellular background signal is very low in this region: i) cellular nucleic acids are mostly extra-large macromolecules, which are not visible by solution NMR; ii) cells do not visibly contain important populations of any small oligonucleotide that would yield detectable imino NMR signals. Importantly, the authors attached a fluorophore in a sub-stoichiometric 40:1 or 50:1 ratio to evaluate the electroporation efficiency (FACS) and the cellular localization of the oligonucleotides (confocal microscopy): they all moved spontaneously to the cell nucleus. Using this approach, Trantirek and colleagues could prove the stability in cells of i-motifs structure adopted by C-rich sequences from the human genome.[631] The ^{1}H-NMR signals in the imino region were indeed observed, which proved that they were engaged in base pairing. They also showed that a number of i-DNA motifs are more stable in cells than in dilute solution.[631,634] The achievable intracellular concentration was enough to detect a RNA-14mer in HeLa cells but not a 72-nucleotide aptamer.[294] Katahira used also one-dimensional ^{1}H-NMR spectroscopy but delivered DNA oligonucleotides using pore-forming toxins. Because of cell leakage issues, they used recently a flow-probe bioreactor, which efficiently removed leaking DNA and maintained cells viable in the NMR spectrometer through extended periods of time. This permitted to record high S/N ^{1}H-NMR spectra through ~16-20 hours of oligonucleotides at ~5-20 µM in ~2-4.10^{7} HeLa cells. They could prove the stability of DNA triplex structures in cells by detecting its ^{1}H-NMR fingerprint (Figure 25).[328]

Xu and colleagues used pore-forming toxins to translocate DNA and RNA oligonucleotides, which were previously tagged with a hexafluorinated probe at the 5' their extremity. This generates ^{19}F-NMR signals at different frequencies in function of the adopted tertiary structures. They could record one-dimensional ^{19}F-NMR spectra of various oligonucleotides at intracellular concentrations of about 150 µM in about 80 minutes. This helped them to show that human telomere DNA sequences could adopt various hybrid-type or two-tetrad G-quadruplex structures in cells. They could also observe the concomitant existence of hybrid DNA:RNA and RNA:RNA quadruplexes.[450] They also introduced a trifluoro-2'deoxyguanosine, which stabilized a Z-DNA conformation *in vitro* and in cells, according to the one-dimensional ^{19}F-NMR fingerprint.[450] In a recent report, Trantirek and colleagues have shown that in-cell ^{1}H-NMR analysis of G-quadruplexes can be hampered when conformational ex

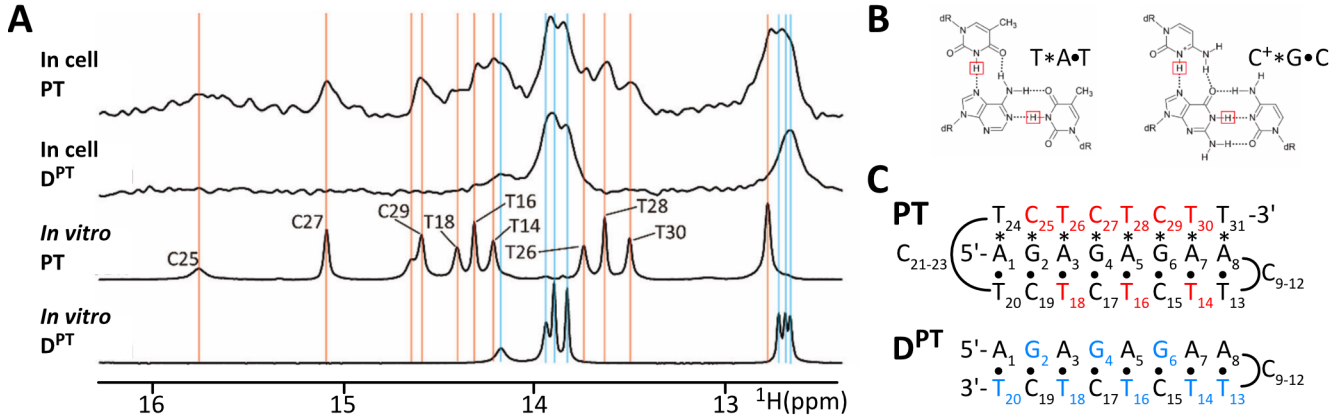


Figure 25: A) One-dimensional ^{1}H-NMR spectra of Parallel Triplex oligodeoxynucleotides (PT) and of their duplex portions (D^{PT}) in vitro and in HeLa cells; the spectrum of PT in-cell reveals the presence of both PT and D^{PT} in proportions that correspond to the appearance of degraded species of PT in cells, according to gel electrophoresis analysis; **B)** Novel H-bond stabilizations of imino protons appearing in the

triplex structure; **C)** Sequences and secondary structures of PT and D^{PT} used in this study (dots and stars indicate Watson-Crick and Hoogsteen base pairs, respectively; red and blue nucleotides correspond to the observed imino protons in PT and D^{PT}). Adapted from Sakamoto et al. ChemComm 2021.[328] Copyright 2021 Royal Society of Chemistry.

change provokes the appearance of too many,and sometimes too broad ^{1}H peaks. DNA integrating a single fluorinated probe can help, even though ^{19}F-NMR generates broad NMR signals in cells: conformation-specific ^{19}F-NMR peaks can overlap and also require deconvolution to be interpreted.[635]

We shall mention recent developments, which take advantage of ^{13}C-^{19}F labeling: this permits to obtain sharper ^{19}F-signals of large proteins and nucleic acids. Although much more expensive, these labeling schemes could find applications for in-cell NMR studies.[118,119,137,139,410] Also, cost-effective protocols exist to produce uniformly or nucleic acid-specific ^{2}H-,^{13}C and/or ^{15}N-labeled nucleic acids.[126–131] Such labeling would permit to record two-dimensional ^{13}C/^{15}N-filtered spectra, which can yield structural information of superior quality.[636] The late Covid-19-related context urged for studies on RNA viruses and RNA-based vaccines. NMR spectroscopy may be helpful in this field to characterize uptake and processing of nucleic acids non-destructively and at atomic-resolution.[637]

3.2. The specific case of disordered proteins

About 40% of the human proteome is made of regions or proteins that do not have a stable fold by themselves.[10,638] These are commonly called "intrinsically disordered proteins" or regions of proteins (IDPs/IDRs). These play fundamental roles in cell signaling: the eukaryotic organisms use notably their capacity to establish low-affinity but high specificity-interactions, which can be reversibly switched on or off by post-translational modifications.[8,639–642] They have impacts at multiple other levels, like the thermodynamic and kinetic modulation of protein interactions, the establishment of allosteric mechanisms, or the organization of membraneles organelles via liquid-liquid phase separation.[643–649] They are also infamous for their aggregates related to a number of protein misfolding diseases.[650,651] NMR spectroscopy had invaluable contributions in the field: it is the only technique likely to provide residue-specific information on these extremely flexible objects.[9,47,652,653]

IDPs have moreover NMR-favorable properties: because of their high flexibility, their resonances relax slower than those of folded proteins, which translates into higher S/N of their NMR spectra. Exploitable 2D ^{1}H-^{15}N spectra of IDPs in dilute solution can be recorded at 1-5 μM in about one hour.[428] The quality of their in-cell NMR spectra is often less affected by the intracellular unspecific interactions and tumbling deceleration, as compared to what is observed with folded proteins. Hence, protein constructs integrating both disordered and folded domains generated NMR spectra in *E. coli*, where only the disordered fraction is visible.[156,180] IDPs interactions have indeed only local impacts on the protein mobility: only the binding segment is engaged and the rest of the protein remains independently flexible and NMR-visible, in most cases.

IDPs spectra are generally poorly dispersed: every residue can explore its local conformational space with a high degree of freedom, which averages out the perceived chemical environments and thus provokes poorly differentiated chemical shifts. NMR spectra of IDPs are thus rather straightforward to recognize, as compared to the more scattered ones of folded proteins. This feature has been used to confirm the disordered state of a number of ^{15}N-labeled overexpressed proteins in *E. coli* cells. The model IDP α-synuclein was particularly scrutinized by four independent groups, using 2D ^{1}H-^{15}N NMR spectroscopy,[154,156,166,174,175,177] or even ^{19}F-NMR.[513] These observations permitted notably to disprove allegations about the folded nature of recombinant α-synuclein when gently purified from *E. coli* cells:[654] its in-cell NMR spectra are those of an IDP before any lysis or purification step.[177] Waudby and colleagues demonstrated it further, by recording in-cell 3D ^{1}H-^{13}C-^{15}N NMR spectra to obtain ^{13}C-backbone chemical shifts, which did not show any signs of secondary structure.[175] Using ^{1}H-^{15}N NMR, Shekthman and colleagues showed that the Mycobacterium prokaryotic ubiquitin-like protein (Pup) was disordered upon recombinant expression in *E. coli* cells.[185] Similarly, Hough and colleagues have shown that FG-rich repeats from yeast nucleoporins (FG-Nups) were disordered upon recombinant overexpression both in *E. coli* and in yeast.[179,426] The same two-dimensional ^{1}H-^{15}N NMR experiments revealed a disordered conformational behaviors for Tau and for the N-terminal residues of the kinase c-Src previously microinjected in frog oocytes.[283,284]

The next step came with the experimental proof that the structural disorder of human disordered proteins persisted in cultured human cells. This was indeed demonstrated by three independent groups, who reported 2D ^{1}H-^{15}N NMR spectra of ^{15}N-labeled α-synuclein in a number of cell lines (including neuronal ones), either after electroporation delivery or upon *in situ* overexpression.[176,270,360] These spectra were indeed superimposable to those measured in dilute solution, where α-synuclein is the archetypal IDP. Minor differences existed though, which were of high importance. First, Theillet et al. observed that some NMR peaks had weaker intensities in cells, which were corresponding to residues in the vicinity of Phe4 and Tyr39 (Figure 26).[176] The alanine-mutation of these two residues provoked to recover high peak intensities in the two regions. Later, Burmann and colleagues showed that these interactions were mainly due to endogenous chaperones: inhibiting HSP90 restored also the NMR signal intensities of residues neighboring Phe4 and Tyr39.[360] Even more important, upon Hsc70 silencing and HSP90 inhibition, they observed that the 2D ^{1}H-^{15}N NMR spectra of α-synuclein resembled those of lipid-membrane bound α-synuclein; they confirmed α-synuclein colocalization with mitochondria using confocal microscopy. This was accompanied by the formation of α-synuclein aggregates. These NMR experiments were carried out on samples made from recombinant ^{15}N-labeled α-synuclein electroporated at intracellular concentrations ranging from 3 to 100 μM in 5-15.10^{7} cells. One hour-long acquisitions produce exploitable spectra at 10-20 μM concentrations.

Interestingly, supplementary interactions were detected on α-synuclein C-terminal acidic regions, which were attributed to unknown cellular cationic species.[176] What is the impact of all these interactions and of cellular viscosity on the global conformational dynamics of α-synuclein? By measuring and analyzing in-cell T1/T2/heteronuclear-NOE relaxation

parameters in dedicated 2D ^{1}H-^{15}N spectra, Theillet al. managed to deconvolute the NMR effects of slower dynamics at the nanosecond timescale from those of interactions at the µs-ms timescale: it appeared that the residue specific correlation times were uniformly about 1.5- to 1.8-fold slower in cells than in dilute solution. Moreover, Theillet et al. measured residue-specific PREs in 2D ^{1}H-^{15}N spectra of in-cell α-synuclein previously cysteine-attached to paramagnetic tags. This produced intramolecular distance profiles very similar in cells and in dilute solution, which revealed only a modest compaction in cells, while a crowding agent like ficoll provoked an apparent compaction *in vitro*.[176] Interestingly, the only region that did not experience any intracellular interactions was the central amyloidogenic region of α-synuclein. The NMR residue-specific information was clearly instrumental there.

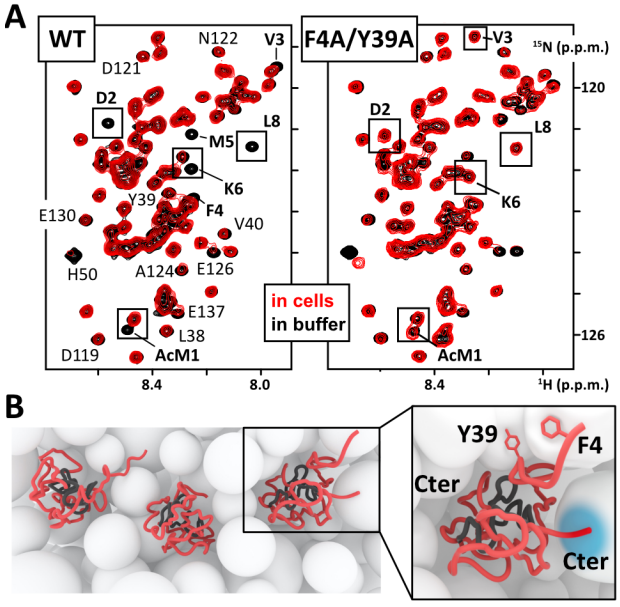


Figure 26: A) 2D ^{1}H-^{15}N LHSQC spectra of ^{15}N-labeled α-synuclein in cultured mammalian cells (A2780 cell lines; spectra obtained in HeLa, and in the neuronal B65, SK-N-SH and RCSN-3 cell lines were about the same).[176] The protein was expressed recombinantly in *E. coli*, purified and delivered in mammalian cells using electroporation. The protein was N-terminal acetylated after its delivery, as shown by the NMR signals of the nine first N-terminal residues adopting frequencies specific of this form. The spectrum of the wild-type protein shows peak intensity losses in the N-terminal, in the vicinity of Y39 and in the C-terminal; the spectrum of the F4A-Y39A mutant revealed a recovery of peak intensities in the N-terminal regions, but not in the C-terminal (the chemical shift perturbations due to F4A-Y39A mutations are particularly visible on Val3, Met5, L38 and V40). **B)** Cartoon of the conformational behavior of α-synuclein in cultured mammalian cells, as interpreted from in-cell NMR and EPR information. The protein is disordered, protects its central NAC region free of any interactions and experiences transient interactions in the regions surrounding Phe4, Tyr39 and its acidic C-terminus. Burmann et al. showed that cellular chaperones are main contributors to the interactions with Phe4 and Tyr39.[360] Adapted from ref [176]. Copyright 2016 Springer Nature.

The nature of IDPs' NMR spectra changes drastically upon binding to folded proteins: peaks can appear out of the spectral region, where disordered residues signals fall normally. However, IDPs experience often loose interactions, and the exchange between bound and free states occurs commonly at the µs-ms timescale, which provokes NMR peak intensity losses. An ambiguity can remain in the cause of NMR peak disappearance in cells: unspecific, transient interactions or folding accompanied by slow tumbling can produce this effect. All shades of grey exist probably in between these two pure situations. An early report of Pielak and colleagues concluded that the transcriptional regulator FlgM overexpressed in *E. coli* cells adopted a folded state because of some NMR peaks disappearance. This might however be explained by intracellular interactions as well. In a more recent study, they measured H/D exchange rates on backbone amides, which were consistent with a disordered behavior.[162] Recently, Zhang et al. released in-cell spectra of ^{15}N-labeled Tau electroporated in HEK cells, which permitted the same peak intensity mapping: it was basically the same than Tau mixed to microtubules *in vitro*.[367] This corroborates prior knowledge on interactions between Tau and microtubules in cells. Hough and colleagues observed peak disappearance in FSFG motifs of FG-Nups both in *E. coli* and in yeast. The high accessibility of disordered residues makes them probably favored targets of unspecific interactions with all sorts of hydrophobic surfaces in cells.[179,426]

Importantly, we must stress the fact that all these in-cell studies on IDPs have been carried out i) using cell pellets with no nutrient replenishment and ii) at low temperature (about 10 ºC). Indeed, ^{1}H-^{15}N NMR spectra are commonly used because they can provide residue specific information and satisfying sensitivity. This is compromised at high temperature or pH, because the water:amide proton exchange rates become faster. ^{1}H-^{15}N NMR peak intensities are inversely proportional to these rates, and ^{1}H-^{15}N spectra are no longer exploitable above pH7 and 25 ºC usually. Low temperature makes the absence of fresh medium supplementation less problematic, by avoiding cell death. Cells are not in their most native conditions. Although not perfect, these conditions were the only ones available to achieve the reported pioneering studies. The future investigations will certainly take advantage of flow-probes bioreactors. As mentioned above, physiological temperatures and pH will however hamper the classical use of ^{1}H-^{15}N NMR spectroscopy. ^{13}C-detected NMR appears to be the next best possibility,[178] which suffers of lower intrinsic sensitivity. Improved pulse sequences and ^{13}C-optimized probes give access to exploitable NMR spectra of disordered proteins at ~25 µM in about one hour at 37 ºC.[436] Whether they would yield such performances in cells remains to be tested.

We mention briefly that the drastic difference between 2D ^{1}H-^{15}N NMR spectra of folded and disordered proteins has also been useful in studies focusing on folded/unfolded equilibria (see chapters 3.1.2. and 3.4.).

3.3. Chemical modifications

NMR spectroscopy reports for any changes in the chemical environment of an amino acid/nucleotide via its chemical shifts. Hence, NMR became an exquisite technique to study post-translational modifications (PTMs) in a residue-specific and time-resolved fashion,[400,405,407,655] among which phosphorylation,[406,436,656–660] acetylation,[176,661–663] methylation,[664,665] oxidation,[331,666] or Asn/Asp/Gln/Glu degradation.[667–669] Barraud and colleagues introduced recently the

NMR identification and monitoring of RNA post-transcriptional modifications.[299]

IDPs are the main targets of cell signaling-related, reversible PTMs.[13] Consistently, in-cell NMR studies reported mostly PTMs on IDPs, using 2D ^{1}H-^{15}N NMR spectroscopy. ^{15}N-labeled Tau and the disordered N-terminal of c-Src were phosphorylated after being microinjected in frog oocytes.[283,284] Selenko and colleagues managed to monitor the phosphorylation of ^{15}N-labeled a viral SV40 peptide: by recording time series of 2D ^{1}H-^{15}N NMR, the progressive appearance of NMR peaks was observed corresponding to SV40(pS112) followed by SV40(pS111). They also exposed oocytes to progesterone, which triggered Cdk1 activity and the phosphorylation of SV40(T124) in addition to those of S112 and S111; all these phospho-events were monitored simultaneously using 2D ^{1}H-^{15}N NMR.

In contrast, pre-phosphorylated Tau was observed to be dephosphorylated after its electroporation delivery in culture mammalian cells.[367] α-synuclein was N-ter acetylated post-translationally in various mammalian cell lines: the 2D ^{1}H-^{15}N NMR spectra were superimposable to those obtained *in vitro* upon incubation with the N-terminal acetyltransferase NatB.[176] About 80% of the human proteome is N-ter acetylated,[670,671] but α-synuclein was believed to be exclusively co-translationally modified by NatB; these in-cell NMR studies proved that it was also readily processed post-translationally. Methionine-oxidized ^{15}N-labeled α-synuclein was also delivered in mammalian cells, where it was reduced selectively on Met1 and Met4, but not on Met116 and Met127.[331] Finally, the progressive proteolytic processing of ^{15}N-labeled α-synuclein delivered in mammalian cells was also characterized using 2D ^{1}H-^{15}N NMR. All these studies require good references and peak assignments *in vitro*: the PTMs are identified by comparing the spectroscopic fingerprints of in-cell spectra and their *in vitro* counterparts.

Cysteine redox equilibrium has also been investigated using in-cell NMR. Cysteines chemical shifts are naturally depending on the oxidation state of their side chain, i.e. thiol or disulfide. Banci and colleagues characterized SOD1 cysteine side chains in cultured mammalian cells overexpressing this protein in a medium supplemented with ^{15}N-cysteine.[237,238,672] They had also used this labeling strategy in a previous study in *E.coli* cells.[171] Hence, they could test the consequences of co-chaperone expression or copper supplementation on the four cysteines of SOD1 in a residue-specific fashion. However, cysteine thiol-disulfide exchange has often an important impact on the neighboring amino acids, or even on the global fold of a protein. Banci and colleagues have thus carried out many other studies using a uniform ^{15}N-labeling obtained by supplementing the mammalian cell culture medium with ^{15}N-amino acids. Using uniform ^{15}N-labeling and ^{15}N-cysteine labeling, they could show that Mia40 and Cox17 were maintained in a reduced state by glutaredoxin 1 and thioredoxin 1.[172,673] These two labeling schemes were also useful to show that Cd^{2+} exposure provoked the oxidation of SOD1 cysteines.[259] Shimada and colleagues adopted another strategy: they produced ^{13}C-Alanine labeled thioredoxin (Trx) recombinantly in *E. coli*, which they later delivered in culture mammalian cells using pore-forming toxins. Indeed, Trx-Ala29 generates a well-separated ^{1}H-^{13}C methyl NMR signal in 2D ^{1}H-^{13}C NMR spectra, which is sensitive to the establishment of a disulfide bond between Trx-Cys32 and -Cys35. Hence, they could monitor the redox state of intracellular Trx upon exposure to the oxidant tert-butyl hydroperoxide (TBH).[326] The authors measured the cellular redox potential simultaneously: they quantified the ^{13}C-labeled glutathione redox equilibrium (GSH/GSSG), which was obtained via ^{13}C-cysteine supplementation and incorporation in endogenous glutathione. Interestingly, the Trx redox transition did occur at a less reducing potential in cells than *in vitro* (Figure 27).

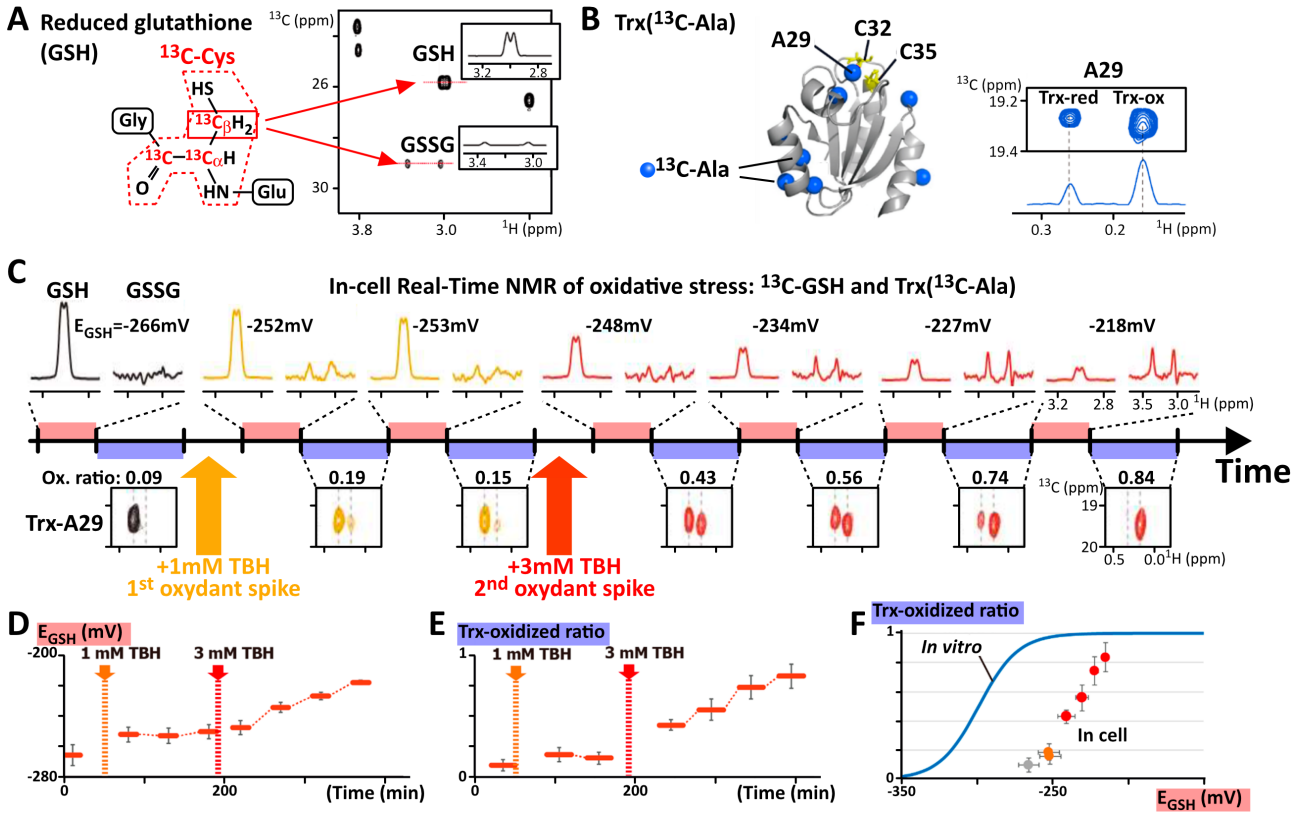


Figure 27: Operating principle of simultaneous in-cell NMR monitoring of the redox potential via GSH/GSSG and of the Thioredoxin (Trx) redox equilibrium. **A)** ^{13}C-Cysteine is integrated by cultured mammalian cells into glutathione, which enables the quantification of GSH and GSSG via the peak intensities of ^{1}H-^{13}Cβ in 2D ^{1}H-^{13}C CT-HSQC spectra; the insets show the peak intensity profile along the ^{1}H-dimension; **B)** Trx(^{13}C-Ala) was produced recombinantly in *E. coli*, purified and translocated in cultured mammalian cells (3.10^7 HeLa cells) using pore-forming toxins resealed via Ca^{2+} supplementation; ^{1}H-^{13}Cβ NMR signal from Ala29 is a good reporter of the redox equilibrium between reduced Cys32/C35 and disulfide bonded Cys32-Cys35; here the 2D spectra are acquired using a ^{1}H-^{13}C SOFAST-HMQC pulse sequence for a better S/N, which cannot be used for ^{13}C-glutathione; **C)** Experimental scenario of an in-cell time-resolved monitoring experiment of the GSH/GSSG and Trx-redox equilibria: 2D ^{1}H-^{13}C CT-HSQC (red periods) and ^{1}H-^{13}C SOFAST-HMQC (blue periods) spectra are acquired alternatively to detect glutathione and Trx, respectively; **D)** Intracellular redox potential in function of time, as measured by the GSH/GSSG ratio in 2D ^{1}H-^{13}C spectra; **E)** Intracellular Trx-redox ratio in function of time, as measured via the Ala29 peak intensities 2D ^{1}H-^{13}C spectra; **F)** Intracellular Trx-redox ratio in function of intracellular redox potential; the blue curve shows the same Trx-redox ratio as measured in dilute solution. Adapted from ref [326]. Copyright 2018 American Chemical Society.

Unfortunately, most of these PTMs could not be monitored in real time, because they were already complete when starting the NMR acquisition: it takes from 2 to 5 hours to rest the cells after protein delivery, before settling them in the NMR tube. This limits considerably the capacity of NMR to describe intracellular PTM mechanisms. Cell extracts can represent an interesting alternative, and they have been used fruitfully in a number of cases.[284,298,331,357,661,674–677] However, we can think to many biases that can make cell extracts a misleading material too: many species precipitate upon lysis, colocalization is lost, the PTM-operating enzymes are possibly not regulated anymore, the energy depletion occurs soon… NMR spectroscopy is still very useful *in vitro* to decipher complex PTMs mechanisms with multiple events, which is of great help to design mutants for complementary cell biology investigations.[45,659,677,678]

3.4. Metal chelation

About 30-40% of proteins integrate metal ions in their structure, a percentage that is also found in the PDB.[679] Metals are often found in catalytic centers or can be pure structural elements, sometimes as triggers of conformational changes (e.g. calcium and calmodulin). They are also suspected to play a role in neurodegeneration diseases.[680,681] How do metal proteins exert their metal preferences, undergo metal shrinkage or oxidative stress, or accommodate competing and eventually harmful metals?[14,682,683] Techniques have been developed to characterize metal binding by proteins, notably using mass-spectrometry. However, it relies on destructive processes, which can produce misleading results: the intracellular binding of metals is regulated by a number of delivery pathways and localized equilibria, which are lost upon cell lysis and can generate irrelevant metal release or competitions.[684–686] Obtaining information from intact cells is thus desirable. NMR spectroscopy can report for metal chelation like for any other changes in chemical environment,[266,687,688] including in cells.[689]

Paramagnetic species provoke fast relaxation and low NMR signal, which made them too challenging for extensive in-cell NMR studies until now. (NB: In-cell structural studies using lanthanide tags do not dealt with any biological question about metal chelation, see chapters 2.3.6. and 3.1.1.). A number of proteins chelate the diamagnetic Mg^{2+}, Ca^{2+}, Zn^{2+}, or Cu^{+}, which are exquisite targets for in-cell NMR studies. An early study also reported the complexation of Pt(II) by

the ^{15}N-labeled Cu-chaperone HAH1 (also named Atox1, 68 amino acids) overexpressed in *E. coli* cells.[423] Indeed, its in-cell 2D ^{1}H-^{15}N HSQC spectrum upon bacterial exposure to cisplatin was superimposable to that obtained in dilute solution with purified Atox1 in presence of cisplatin. Hence, the in-cell spectra confirmed that this interaction was also possible in a cellular environment, where many other metal ions might have outcompeted Pt(II) chelation by Atox1.

At the opposite, Banci and colleagues reported examples of metal binding events that were observed *in vitro*, but prevented in cells. The copper, zinc superoxide dismutase SOD1 must chelate one Zn^{2+} and one Cu^{+}/Cu^{2+} to be active, in addition to dimerizing and establishing intramolecular disulfide bonds. Understanding its maturation scenario is thus rather complex and was rather puzzling: for example, it can chelate two Zn^{2+} in its two metal binding sites *in vitro*. Interestingly, this did not occur in *E. coli* cells supplemented solely with Zn^{2+} and deprived of copper: the recombinant, overexpressed ^{15}N-labeled SOD1 generated in-cell 2D ^{1}H-^{15}N spectra that did correspond to E-Zn-SOD1 but not to Zn-Zn-SOD1, according to the *in vitro* reference spectra.[171] However, copper supplementation did not provoke the appearance of the expected NMR spectrum or Cu-Zn-SOD1. This only occurred once SOD1 was co-overexpressed with its cochaperone CCS in cultured mammalian cells. Banci and colleagues recorded 2D ^{1}H-^{15}N NMR spectra of ^{15}N-labeled SOD1 in HEK293 cells supplemented or not with Zn^{2+} and Cu^{2+}, and with and without stoichiometric expression of SOD1: these revealed that only the combination of Zn^{2+}, Cu^{2+} and CCS permitted to obtain the final mature dimer Cu-Zn-SOD1, while the absence of CCS yielded E-Zn-SOD1.[237,238] In contrast, a series of SOD1 mutants from patients with familial amyotrophic lateral sclerosis (fALS) were found to be incapable of chelating Zn^{2+} in absence of CCS, as revealed by their in-cell 2D ^{1}H-^{15}N NMR spectra.[255] Surprisingly, the co-expression of CCS provoked their correct folding into Cu-Zn-SOD1, notably by stabilizing the E-Zn-SOD1 intermediate.[257] Banci and colleagues showed also that Cd^{2+} was unable to replace copper in cultured mammalian cells, which was revealed by the only observation of NMR spectra of E-Zn-SOD1.[259] Using the same ^{15}N-labeled expression in HEK293T cells and in-cell 2D ^{1}H-^{15}N NMR spectroscopy, they also demonstrated that the 20 kDa Parkinson's disease related DJ-1 did not bind zinc or copper in cells, while it did in dilute solution. We shall mention the authors also showed that metal chelation by SOD1 was also detected via one-dimensional ^{1}H-NMR spectroscopy, a faster and more sensitive approach: the histidine side chains resonate in background-free region of the ^{1}H spectra (about 12-16 ppm), and, being involved in metal chelation, their chemical shifts are specific reporters of this interaction. They do not provide the rich information about SOD1 folding and redox equilibrium that 2D ^{1}H-^{15}N spectra reveal though.

Ca^{2+} is extremely important for cell signaling and can trigger a number of protein structural and activity changes in cells.[690,691] Calmodulin (CaM) may be the most famous Ca^{2+} sensor, whose Ca^{2+}-bound conformation permits to interact and regulate multiple enzymes.[692] CaM undergoes a conformational rearrangement upon accommodating two Ca^{2+} ions at two chelation sites, which might be affected by cellular crowding and the interactions with its numerous protein partners. Hence, CaM has been a target of in-cell NMR since the beginnings, and it has been scrutinized in a number of studies.[147,148,167,168,189,192,193,226,227,280,286] However, it yields often poor spectra, either because of its intrinsic capacity to exchange between its different conformations, or to interact with too many cellular entities. Li and colleagues circumvented this problem by labeling CaM with unnatural fluorinated 3F-Tyr. They designed a mutant of CaM with only two tyrosine residues generating two peaks in 1D ^{19}F-spectra, whose chemical shifts are different in the Ca^{2+}-free and Ca^{2+}-bound forms. These chemical shifts are further perturbed upon interacting with Ca^{2+}-activated CaM-binding peptides, like a fragment of the myosin light-chain kinase (MLCK). Interestingly, Li and colleagues introduced preformed Ca^{2+}-CaM at 0.22 mM in frog oocytes, but observed ^{19}F-NMR spectra corresponding to the apo-CaM.[192] While the Kd of Ca^{2+} binding is about 7.10^{-6} M in vitro, they had to increase the intracellular [Ca^{2+}] up to 2 mM to reach an equimolar ratio apo-CaM:Ca^{2+}-CaM. The microinjection of Ca^{2+}-CaM:MLCK complex at 0.22 mM revealed a 1:1 ratio between apo-CaM and the complex Ca^{2+}-CaM:MLCK in frog oocytes. Ca^{2+} ions were probably involved in other competing events in cells, or even scavenged or pumped out, and MLCK binding stabilized Ca^{2+} binding in cells as it does *in vitro*.

In a short study, Ito and colleagues have introduced ^{15}N-labeled calbindin D_{9K} in cultured mammalian cells (HeLa) using the cell-penetrating peptide approach.[312] This protein (79 residues) acts as a calcium-buffer in the cytosol, where it binds Mg^{2+} in absence of Ca^{2+} (affinities are separated by 1000-fold). While they introduced the Ca^{2+}-loaded form of D_{9K}, they recorded in-cell 2D ^{1}H-^{15}N NMR spectra corresponding to the Mg^{2+}-bound form initially. However, their sample was simply made of a cell pellet, without any fresh medium replenishment. After 2 hours at 37 ºC in the NMR tube, it generated 2D ^{1}H-^{15}N NMR spectra corresponding to Ca^{2+}-bound D_{9K}, consistent with a loss of cell homeostasis and a release of intracellular Ca^{2+}.[312]

We must notice that these NMR spectra were recorded at intracellular concentrations of the studied proteins between 50 and 300 μM.[172,237,238,255–261] The case of SOD1 and the necessary co-expression of its cochaperone CCS remind us that a good prior knowledge of the biological system is important. The functional partners of the studied protein should better be expressed at similar levels. In any case, we have seen that purified material or from cell lysates can produce misleading observation in this field. To the best of our knowledge, in-cell NMR studies represent the only approach to obtain such a relevant and detailed information on metal binding by proteins.

3.5. Protein:protein specific interactions

NMR spectroscopy is a method of choice to describe macromolecular interactions at the residue scale.[47,693,694] We have discussed oligonucleotides structures in the chapter 3.1.3, which often depend on their interactions; we will thus not evoke them further in the present chapter.

In early studies, Shekhtman and colleagues showed that protein:protein interaction surfaces could be mapped in *E. coli*, by expressing successively a first ^{15}N-labeled protein and a second or even a third one at ^{14}N-natural abundance.[149,182] We described the protocols in chapter 2.2.2., and its yeast counterpart in chapter 2.2.3. This permits to obtain 2D ^{1}H-^{15}N NMR spectra of only one protein, hence avoiding spectral crowding and difficulties or ambiguities in the later

residue-peak specific analysis. They mapped the in-cell interactions on ^{15}N-labeled surface with the ubiquitin-interacting motif from ataxin 3 (AUIM) or with the signal-transducing adaptor protein 2 (STAM 2), based on chemical shift perturbations or peak broadening.[149,182] Only minor differences were observed with the same interactions analyzed *in vitro*. Then, they showed that ^{15}N-ubiquitin interaction surface with STAM2 was affected upon co-expression of the tyrosine kinase Fyn in *E. coli* (STAM2 and Fyn were non-labeled).[183] The interaction surface with the Hepatocyte growth factor regulated substrate (Hrs) was not affected upon Fyn expression, suggesting that STAM2 phosphorylation was responsible for the pervious variation. Shekhtman and colleagues used the same approach to detect and map the interaction between the ^{15}N-labeled fujimycin binding protein (FKBP) and the non-labeled rapamycin binding protein (FRB) in *E. coli*.[184] They exposed the cells to known FKBP-binding inhibitors of this interaction, which provoked chemical shift changes in the in-cell 2D ^{1}H-^{15}N NMR spectra of FKBP, consistent with prior knowledge. Finally, they overexpressed the *Mycobacterium tuberculosis* prokaryotic ubitin-like protein (Pup) in a ^{15}N-labeled fashion, and later expressed the mycobacterial proteasomal ATPase (Mpa),[185] or identified Pup-binding peptides,[186] in a ^{14}N-labeled fashion in *E. coli*. The in-cell 2D ^{1}H-^{15}N NMR spectra of ^{15}N-Pup showed residue-specific peak broadening in presence of Mpa or of the Pup-binding peptides, which allowed a residue-specific mapping, and eventually a drug screening.[187]

Shimada and colleagues used in-cell NMR to verify a weak interaction characterized previously *in vitro*: the 9 kDa domain CG1 of the microtubule-binding protein CLIP-170 (CG1) interacts transiently with acidic C-terminal peptides of tubulins with a low affinity (Kd~70 µM).[695] The highly electrostatic nature of this interaction might have been too degenerate to be relevant in cells. They delivered deuterated (^{1}H/^{13}C-Ile-Leu-Val)-CG1 at ~180 µM in HeLa cells using pore-forming toxins, and recorded 2D ^{1}H-^{13}C NMR spectra, which gave access to residue specific information.[267] The observed in-cell chemical shifts were the same than those observed *in vitro* in presence of purified microtubules. Moreover, proton spin saturation (via radiofrequency irradiation)[696] of the protonated cellular material was preferentially transferred to the resonances of the residues identified as contacting microtubules *in vitro*. Finally, they introduced a mix of non-labeled and deuterated (^{1}H/^{13}C-Ile-Leu-Val)-CG1 at a ratio 10:1 to provoke an intracellular binding competition, which resulted in the loss of residue specific saturation. This saturation was thus revealing a real intracellular binding event. Such a scheme might be applied to heterodimers, where one protein would be protonated and the other deuterated.

We have just seen in the previous chapter that ^{19}F-calmodulin generated in-cell ^{19}F-NMR spectra once microinjected in frog oocytes. The ^{19}F-Tyr peaks were sensitive to Ca^{2+} binding, so were they to calmodulin-binding peptide MLCK. Hence, Li and colleagues could quantify the stoichiometry of the Ca^{2+}-calmodulin:MLCK complex in cells.[192]

^{19}F-Trp labeling was also used by Pielak and colleagues to evaluate the monomer-dimer equilibrium of a A34F mutant of GB1, which has a Kd of ~40 µM *in vitro*. GB1(A34F) contains only one Trp, and it yields two well separated peaks for the monomer and the dimer forms. The authors overexpressed 6F-Trp-GB1(A34F) (see chapter 2.3.4.) in two types of *E. coli* cells: in a BL21(DE3) strain, only the dimer 19F-signal of 6F-Trp-GB1(A34F) appeared upon IPTG-induction at 25 and 1000 µM; at the opposite, in the Tuner™ strain, the monomer or the dimer peaks were dominant at 25 or 1000 µM IPTG, respectively.[202] The reason is the following: Tuner™ cells, whose lactose permease is deleted, are all equally permeable to IPTG, and express the recombinant protease in an IPTG-concentration manner; in contrast, BL21(DE3) cells are penetrated are induced by IPTG in a stochastic, binary manner, so that individual cells overexpress or not the recombinant protein.[150] Hence, BL21(DE3) cells either contain 0 or ~2 mM 6F-Trp-GB1(A34F), which yield only a dimer peak of variable intensity depending on the IPTG concentration and the number of cells that are induced in the sample. Next, using the Tuner™ strain, Pielak, Li and colleagues expressed 6F-Trp-GB1(A34F) at various intracellular concentrations *E. coli* cells, and quantified the monomer:dimer using one-dimensional in-cell ^{19}F-NMR spectroscopy: they measured a K_d of 11±4 µM in cells versus 37±4 µM in dilute solution (Figure 28).[196]

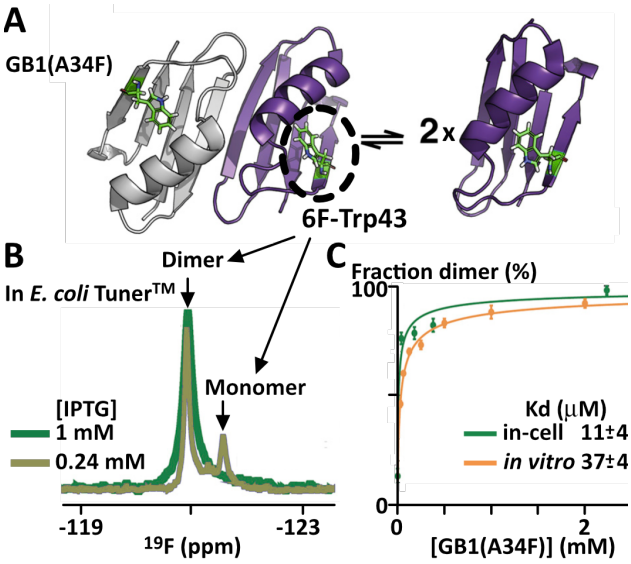


Figure 28: Dimerization of GB1(A34F) as analyzed using fluorinated non-natural amino acids incorporation and ^{19}F-NMR spectroscopy: **A)** Cartoon representation of GB1(A34F) dimerization equilibrium; the Trp43 residue is represented in green sticks; **B)** 6F-Trp-labeled GB1(A34F) expressed in *E. coli* Tuner™ cells generates two peaks at different frequencies for the mono- and di-mer forms; the total concentration of intracellular GB1(A34F) depends on the concentration of IPTG used for the induction of expression; C) Percentage of dimer in the total GB1(A34F) population in function of the total GB1(A34F) concentration in *E. coli* Tuner™ cells (green) and *in vitro* (orange). Adapted from ref [196]. Copyright 2021 National Academy of Sciences.

Obviously, in-cell structural biology by NMR requires rather high intracellular concentrations, above the micromolar range. It can thus apply to interactions of low micromolar affinities, even though competitive assays can theoretically give access to measurements of affinities about two orders of magnitude lower.

3.6. Promiscuous interactions and rotational/translational diffusion

The in-cell NMR community has always been preoccupied by the extent to which the studied proteins engage in loose, non-functional interactions with the cellular components. Although they correspond to short lifetime events, these promiscuous, unspecific interactions can be very deleterious to the NMR signal, by provoking fast T2 relaxation, hence broad linewidth and low S/N. Conversely, evaluating T1 and T2 relaxation can inform about how much a protein interacts with its cellular environment, and what are the leading forces of it. We summarized the techniques used in this purpose in chapter 2.3.7., and we will only focus on the conclusions in the present chapter. Interactions with cellular entities have also been invoked for folding/unfolding equilibria, but we refer the reader to chapter 3.1.2. for these aspects.

Obviously, functional interactions are possible causes of in-cell NMR signals broadening, as shown by Dötsch and coworkers on the proline isomerase Pin1, which is known to interact with a great number of targets.[697] While the apo-Pin1-WT microinjected in frog oocytes was undetectable, they could record exploitable in-cell ^{1}H-^{15}N NMR spectra with a Pin1 inactive mutant, or by saturating Pin1-WT with a genuine binding-peptide (Smad3) before microinjection. These results suggest that Pin1 NMR signals were broadened because of multiple interactions with a wealth of possible intracellular peptides.

However, even in absence of expectable functional interactions, it was noticed very soon that ^{15}N- or ^{13}C-labeled proteins, either overexpressed in *E. coli* cells or microinjected in frog oocytes, were generating spectra of very different qualities. Molecular weight or cellular viscosity were not the main factors, because two proteins of similar size could produce either exploitable spectra or broad, almost undetectable NMR signals.[147,170,173,189,280] Unspecific interactions with cellular entities were immediately proposed to be responsible of the detected NMR line-broadening. "Unspecific" is a matter of definition, but we can safely hypothesize that human ubiquitin does not have any functional partner in *E. coli* cells,[173,280] as well as GFP in frog oocytes.[420] Importantly, Shirakawa and colleagues showed that mutating a surface hydrophobic patch (L8A-I44A-V70A in Ubi3A) permitted to recover very decent in-cell ^{1}H-^{15}N NMR signals in frog oocytes.[280] The same was observed with ^{15}N-labeled ubiquitin overexpressed in *E. coli* cells,[173] or CPP-delivered in cultured mammalian cells.[310] In *E. coli* cell lysate, Ubi3A had amide $^{1}H_N$ NMR linewidths about twice sharper than Ubi-WT; this evaluation was not possible in live cells because Ubi-WT NMR signals were broadened beyond detection.[173] The better behaved model protein GB1 provided interesting results in live *E. coli* cells: its apparent rotational correlation time was about 8 times slower in cells than in dilute solution; nevertheless, 30% of its amide $^{1}H_N$ NMR line widths was due to unspecified interactions with cellular components.[173] This $^{1}H_N$-linewidth derived intracellular correlation time was however not a direct reporter of viscosity: it was integrating the effects of both cellular viscosity, and of transient interactions, which slow down the average, apparent tumbling time. Li, Pielak and colleagues proposed a ^{19}F-NMR approach to deconvolute viscosity and (un)specific interactions.[195,197,199] They showed that GB1 and Ubi-WT were experiencing a 2- to 3-fold higher viscosity in *E. coli* cells than in dilute solution.[195] They also evaluated the impact of intracellular weak interactions on ^{19}F-T2 relaxation: they provoked a supplementary, "apparent viscosity", ~5- to 22-fold that of water.[195]

Different groups tried to dissect the causes of these unspecific, transient interactions. A prominent factor is the electrostatic charge of the studied proteins: mutating surface Arg and Lys into neutral or negative residues lead systematically to improved ^{1}H-^{15}N NMR spectra in *E. coli* and cultured human cells;[256,424] at the opposite, introducing positive solvent-accessible residues provokes adverse effects by slowing down the intracellular molecular tumbling of the studied proteins.[197,201,424] Solvent-exposed aromatic acids and hydrophobic patches are also responsible for cellular interactions.[256,424] Transient interactions centered on exposed aromatic residues are clearly observed in the case of disordered proteins,[176,426] which occur with cellular chaperones to a large extent in the case of α-synuclein.[360] Interestingly, the high negative charge density in the C-terminus of α-synuclein is also involved in transient interactions with cellular entities.[176,360]

In an attempt to better understand how surface residues make a protein prone to intracellular diffuse interactions, Danielsson, Oliveberg and colleagues analyzed >130 surface mutants of three small model proteins in *E. coli* cells.[424] Initially, they interpreted the amide $^{1}H_N$ NMR linewidths in terms of apparent tumbling time, a simple model that do not account for the separate effects of intrinsic cellular viscosity and unspecific interactions. This admirable experimental effort permitted to nail an interesting correlation: the slower apparent tumbling time of a protein in live cells, is linked to its surface negative charge density (Figure 29A). This correlation appears to be modulated by the surface hydrophobicity and the dipole moment of the studied proteins (Figure 29B). It is possible to convert these apparent tumbling times in protein-specific apparent viscosity, which appears to be an exponential function of the protein net charge.[425] This is probably due to the average negative charge of cellular proteomes (Figure 29D): a net-positive protein experiences more and stronger interactions with cellular proteins in a random walk than a net-negative protein. However, these apparent viscosities are 6 times higher in *E. coli* than in cultured mammalian cells, which corresponds to the difference in crowding between these cells. Finally, Danielsson, Oliveberg and colleagues proposed an interesting model to interpret further these in-cell NMR-apparent tumbling times: the amide ^{15}N T1 and T2 values can be interpreted in a simple model where the studied protein is in fast exchange (<μs time scale) between free and bound states; they found out that their model would perform better if these fast, transient interactions occurred with a lognormal size distribution of cellular macromolecules.[366] Interestingly, the fitted distribution is close to that of human cytosolic proteins. These relaxation times are classically measured in 2D spectra to obtain residue-specific information on intramolecular dynamics. Here, the authors measured the global relaxation times of six ^{15}N-labeled, model proteins electroporated in mammalian cells: they performed one-dimensional ^{1}H-detected/^{15}N-filtered NMR experiments, which account for the sum of all residues T1 and T2 relaxation. This model permits to fit the bound populations of the studied proteins: one R/K to E mutation decreased the bound populations of the three model

proteins from 10-12% to 5-6%. Altogether, it is striking to realize that relatively simple physico-chemical parameters seem to control the diffusive, unspecific intracellular interactions that a protein undergo: these are the net charge, surface hydrophobicity, the dipole moment and the rapid binding of ~10% of a protein population to cytosolic proteins. Unspecific binding to ribosomes and RNAs appears to also have important contributions to T2 relaxation of proteins, as suggested by Shekhtman and colleagues.[359,395,422]

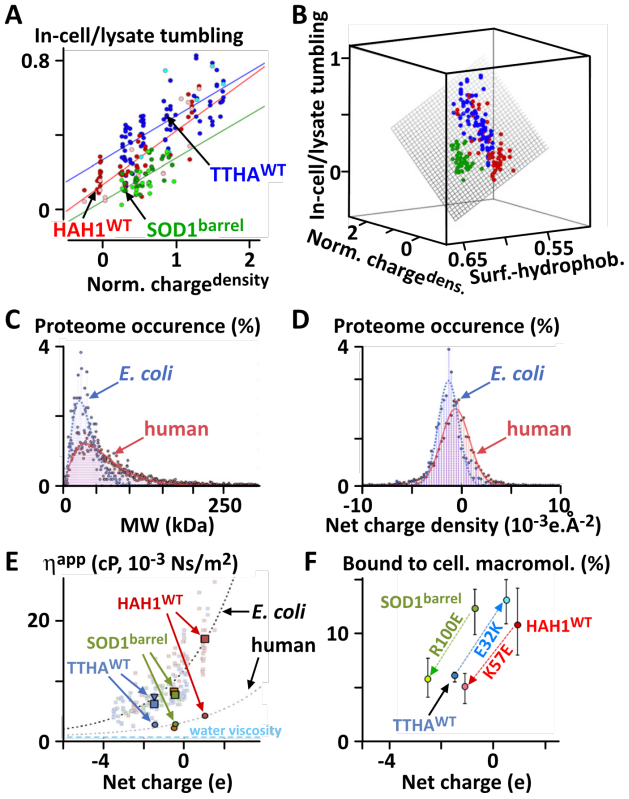


Figure 29: Insights in NMR-apparent mobility of intracellular proteins. The T2 relaxation times of the amide $^{1}H_N$ or ^{15}N are influenced by multiple factors, among which the intrinsic cellular viscosity and the unspecific interactions. Here, we present the works of Danielsson, Oliveberg et al., who did not separate these terms. The aggregated model of in-cell NMR-apparent mobility/tumbling is a direct conversion from $^{1}H_N$ and ^{15}N T2 relaxation. **A)** Ratio of apparent molecular tumbling "in live *E. coli*" over "in *E. coli* lysate" of >130 mutants of the three model proteins HAH1, TTHA and SOD1^{barrel} as a function of the normalized charge density;[424] in-cell tumbling is derived from $^{1}H_N$ linewidth; the normalized charge density was evaluated from the migration distance in native gels, relatively to that of TTHA^{WT}; **B)** Same mobility ratios in function of normalized charge and surface hydrophobicity for the >130 proteins;[424] **C)** Distribution of molecular weights in the cytosolic *E. coli* and human proteomes;[425] **D)** Distribution of surface net charge density in the cytosolic *E. coli* and human proteomes;[425] **E)** Protein-specific, apparent viscosity as a function of net charge in *E. coli* and in cultured human cells; apparent viscosities were derived from global protein ^{15}N T2 relaxation, as measured in one-dimensional ^{1}H-edited/^{15}N-filtered NMR spectra;[425] **F)** Fitted fractions of a protein population binding to cellular macromolecules, as derived from the global ^{15}N T1 and T2 relaxations measured in cultured mammalian cells in the model of fast, random interactions of Danielsson, Oliveberg et al.[366] **A** and **B** are adapted from ref [424], copyright 2017 National Academy of Sciences; **C**, **D** and **E** are adapted from ref [425], copyright 2020 from Leeb et al. under the under the terms of a Creative Commons CC BY license https://creativecommons.org/licenses/by/4.0/; **F** is adapted from ref [366], copyright 2020 American Chemical Society.

NMR spectroscopy is historically an exquisite technique to measure molecular translational diffusion, using the so-called Pulsed Field-Gradient (PFG) NMR.[500–502] It is able to report for rms motions from 10 nm to 100 µm that occur in ~10 to 500 milliseconds However, in the cellular context, proteins show restricted motion above the millisecond time scale, either because of cellular boundaries (~1 µm for bacteria) or because of the subcellular meshwork of cytoskeleton and organelles. Hence measuring rms displacements in 10 millisecond or more is most often poorly informative. To the best of our knowledge, the only NMR-measured translational diffusion of a protein came from Waudby et al., who reported an apparent rms displacement of ~0.3 µm for a-synuclein in E. coli cells for all diffusion delays ranging between 10 and 20 ms: this corresponds to the calculated Brownian motion in a 0.5x0.5x2 compartment, i.e. the size of *E. coli* cells.[174] PFG NMR has been more informative for small metabolites in cell samples, for example, but this is beyond our scope (see Theillet et al. Prog NMR Spectr. 2022 [62]).

Altogether, while NMR spectroscopy has been a method of choice to measure translational and rotational diffusion both in dilute and complex solutions, it has been less productive for measuring macromolecules mobility in cells. Translational diffusion measurements are difficult using the common PFG techniques, which require milliseconds-long diffusion delays, a timescale at which macromolecules displacement is restrained by the subcellular cytoskeleton and organelles. Rotational diffusion is accessible via T1 and T2 measurements, which are also affected by all interactions in cells, specific and unspecific. Deconvoluting these effects is now possible if ^{19}F- or ^{15}N-relaxation times of rigid residues are measured, typically in the hydrophobic core of a protein. Their chemical environment and flexibility do not change drastically during the interactions, and so the effects of pure cellular viscosity and (un)specific interactions can be deconvoluted from the analysis of T1 and T2 values. This gives access to the quantification of specific and unspecific interactions of a protein in cells. We must stress the fact that such quantification is qualitative and linked to interpretation models. It is however very useful to better think the intracellular behavior of macromolecules. It has already been used to produce some renewed representations of the importance of the diffusive protein interactions in cells and how they are affected by basic physico-chemical parameters.[366,425,698] It underlines the importance of the relationship between a protein and the whole cellular environment. It determines the diffusion capacities of a protein, the rate and strength of its cellular interactions and has certainly functional consequences. These may thus be the results of natural selection mechanisms, a reason why some authors tend to see quinary structures and interactions in the observed degrees of intracellular mobility restriction. We have discussed the use of quinary structure in chapter 3.1.2. We would formulate the same cautious note here: quinary interactions would correspond to the selected effects of the relationship between a protein and its native cellular environment. The studied heterologous proteins at high concentrations in cells are not

observed in a native context; we prefer to talk about prefer to use terms like promiscuous or diffusive cellular interactions, for example.

3.7. Questions about protein/nucleic acid localization, concentration, and cellular homeostasis

The endogenous protein concentrations are broadly distributed, averaging in the 0.1 µM range, with only 10% of proteins reaching micromolar concentrations (Figure 30).[18,20] The concentration of a given protein is potentially highly variable through cell-cycle or upon specific cellular events.[699–703] These mean concentrations are actually poorly informative: Proteins/nucleic acids are not evenly distributed in cells, and local concentrations can reach hundreds of micromolar concentrations in organelles and even millimolar concentrations in phase-separated condensates.[23,32,33,703–708] This being said, the average endogenous concentrations are below the standard sensitivity in current protein solution-NMR spectroscopy. To give a rough order of magnitude, 1-hour-long experiments at 600 MHz can yield exploitable 2D spectra of a 30 kDa protein at ~25 µM, in absence of conformational- or binding-exchange. These numbers are of course size- and spectrometer-dependent: the smaller the protein or the larger the field, the higher the NMR signal. In this regard, it is interesting to note that proteins with high copy numbers tend to get smaller, the average size of those found at micromolar being about 30 kDa. [18] Unfortunately, the NMR-effective concentrations are even lower once cells are settled in a NMR tube, where cells can not represent more than 20-40% of the NMR active volume even in wet pellets.

Another important aspect to consider is protein/nucleic acid subcellular localization and detectability. The in-cell NMR signals of a given species report for the average resonances of NMR-detectable molecules, i.e. those i) tumbling sufficiently fast, ii) not interacting with too many intracellular components (see Chapter 2.1.3.). Conformational or binding exchange in the µs-ms timescale can also be deleterious for the in-cell NMR detectability (see Chapters 2.3.7. and 3.6.). Hence, the visible signals are not necessarily accounting for all the molecules, and immediate interpretation might be biased.

Most in-cell NMR studies have been carried out using intracellular concentrations of proteins/nucleic acids in the range of 5 to 100 µM. These high levels can have saturating effects on physiologically relevant interactions, unless the native binding partners are also leveled up. Unfortunately, the observed free molecules would generally yield higher NMR signals than the bound ones, which could also lead to misinterpretations.

Hence, functional cell biology is difficult to approach using in-cell structural biology by NMR spectroscopy. This calls for cautious interpretation. It is notably important to compare these results with the total intracellular amounts of the molecule of interest (using western-blotting, or mass-spectrometry or other methods), and its localization (using microscopies). The main challenge in the next year will certainly be to record NMR spectra at ~1 µM, which is arguably a realistic goal in our opinion (see Chapter 5.2.)

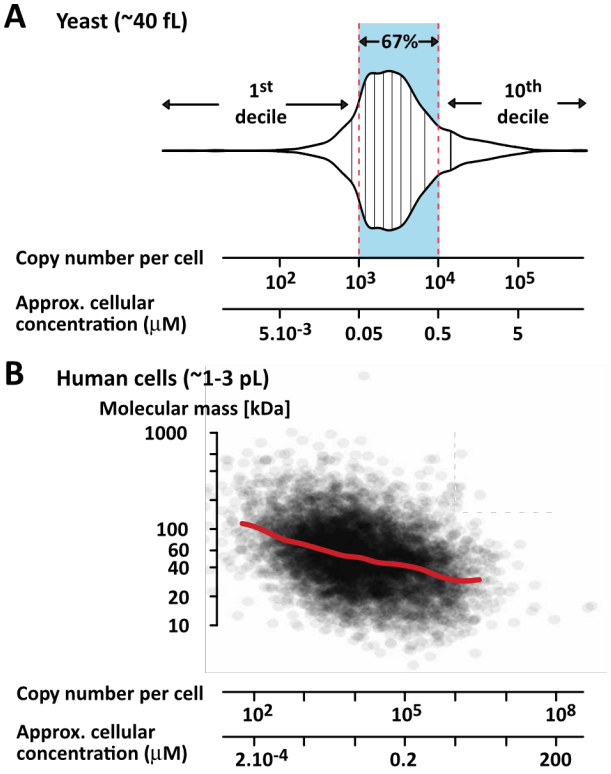


Figure 30: A) Violin plot showing the distribution of protein abundance in yeast cells; vertical lines separate the deciles of the distribution; cellular concentrations were calculated according to an average cell volume of 40 fL; adapted from ref [20], copyright 2018 Elsevier. **B)** Protein size in function of copy number in mouse lung cells; cellular concentrations were calculated according to an average cell volume of 1 pL; adapted from ref [18], copyright 2014 from Wisniezski et al. nder the under the terms of a Creative Commons CC BY license.

4- WHAT BENEFITS FOR THERAPEUTIC PURPOSES?

Structural biology finds important applications in the field of drug research. NMR spectroscopy is an important technique in fragment-based drug-design,[709–711] or in structure-activity relationship studies.[537,711,712] In-cell solution NMR studies have also interesting capacities to help in characterizing drug binding or drug engagement in cells, as we will show in this chapter. SsNMR spectroscopy has had important contributions while using cellular samples or membrane extracts: it permitted to understand how some antibiotics interfere with bacterial cell-wall, either by describing their binding structures, or by describing the changes in peptidoglycan composition (among others, see [64–66] [62]). We restrict our scope to solution NMR here.

In-cell NMR investigations can deliver information linked to drug research topics by characterizing i) interactions via ligand-observed methods, ii) interactions via macromolecule-observed methods, or iii) enzymatic or metabolic activity via the observation of activity-reporters or metabolites (Figure 31). The small molecules can be detected with or without isotope-labeling. Indeed, because of their high concentration and sharp NMR peaks, small compounds generate outstanding ^{1}H-NMR signals that are straightforward to identify and monitor, even in presence of the cellular back-

ground. At the opposite, it is often necessary to isotope-label macromolecules to benefit of the isotope-filter and enable their detection in cellular samples.

Let us underline three important aspects of in-cell NMR applied to drug research: i) it can deliver information on proteins that would be difficult to purify; ii) it is non-destructive and might also be applied on organoids; ii) it permits the parallel monitoring both the small compounds and the protein/nucleic acid targets, of their binding and chemical modifications and even to inform on metabolic outcomes.

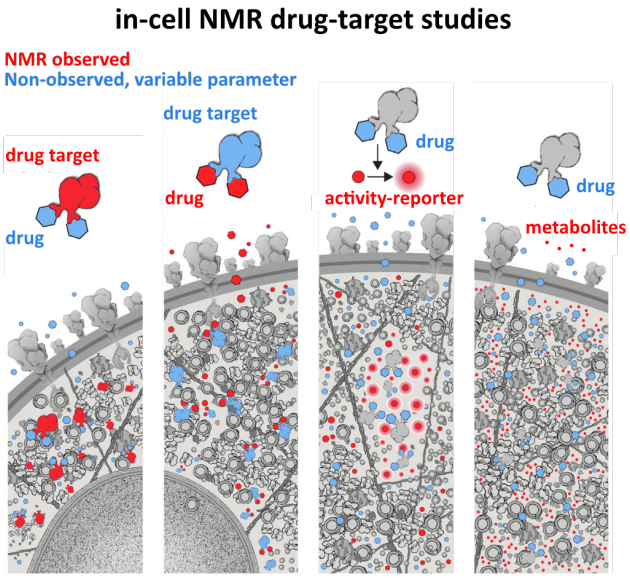


Figure 31: Cartoon representation of the different types of in-cell/on-cell NMR investigations on drugs. The different strategies are classified by their readout: the NMR-observed molecule (red) can be the target-macromolecule, the ligand, an activity-reporter or the cellular metabolites; among others, the varying parameter (blue) can be the exposure to a drug/ligand or the expression/mutation of the macromolecule. Adapted from ref [56]. Copyright 2019 from Siegal et al. under the terms of a Creative Commons CC BY license.

4.1. On-cell NMR and structure-activity relationships

In-cell NMR studies on drug-target interactions focused mostly on membrane proteins. In fact, these might be called "on-cell NMR studies". They have used famous techniques in the field of NMR drug research, namely STD-NMR and TRNOESY (see chapter 2.3.11.). These are typically efficient for middle- to low-affinity interactions, whose Kds range between 10^{-7} and 10^{-3} M and k_{off} between 10^{3} and 10^{6} s^{-1}. The studies evoked below required about 5-10.10^{6} cells in 500 µL, ligands at 0.1 to 3 mM and 1 to 5 hours of acquisition times. They come with a great advantage: they permit to characterize drug binding on membranes proteins without any purification steps![713]

From 2005, pioneering studies could show that STD-NMR was feasible using live cells expressing high levels of membrane receptors, and that this native environment was likely to produce better affinity for cognate ligands.[551,555] Soon after, STD-NMR permitted to verify that a compound identified from screening assays was binding to a cannabinoid receptor CB1 and CB2. STD signals were indeed obtained only in presence of insect cells only when they expressed these receptors.[552] On-cell STD-NMR revealed also the binding of thiol-containing compounds to odorant receptors only in presence of copper and not of other metal salts, in a very complementary fashion to luciferase functional cellular assays.[554,556] STD-NMR has also the capacity to provide information on the binding epitope of a ligand. Haselhorst and colleagues used this feature to progressively improve the chemical structure of compounds interacting with the B-cell receptor Siglec-2 (CD22) and to finally obtain high-affinity ligands.[558] In a similar fashion, Carotenuto and colleagues used on-cell STD information to carry out docking simulations of peptide antagonists of the GPCR CXCR4, and to rationalize the observed differences in affinity.[563] A recent STD epitope mapping has been also released for a native peptide ligand of the GPCR Gastrin Releasing Peptide Receptor (GRP-R) in stably transfected HEK-293T cells.[567] Another interesting application was reported by Potenza and colleagues: they generated stable transfected HEK-293T cell lines expressing influenza hemagglutinin H1 or H5 (human and avian variants, respectively), and compared the STD signals obtained on human/avian-derived cell-surface trisaccharides.[557,714] They could map the epitopes and rationalize the binding properties of the two influenza strains. This approach represents an appealing improvement, as compared to the more classical use of recombinant, solubilized hemagglutinin constructs lacking their transmembrane domains, which may affect their structure and binding properties.

STD-NMR can also be used to obtain direct experimental evidences of drug binding in live cells. This rationale was followed by Izadi-Pruneyre and colleagues, who tested an interaction between a drug target of *Mycobacterium tuberculosis* and a candidate drug recently discovered. They detected STD effects on this drug only in presence bacterial cells expressing the hypothesized wild-type protein target, but not for those expressing a resistant mutant.[547] This target protein, a subunit of a membrane cytochrome supercomplex, was particularly difficult to purify. STD effects confirmed also the binding of a peptide to the vascular endothelial growth factor receptor 2 (VEGFR2),[564] and of a drug candidate to Bcl-2 in live cells.[565]

TRNOESY has been a good complementary technique for on-cell NMR studies. It yields negative of positive NMR peaks in 2D ^{1}H-^{1}H spectra for small ligands, which bind or do not bind macromolecules, respectively.[546,571,572] It is usually less sensitive than STD-NMR when k_{off} decreases, but it can provide structural information on the bound form of the ligand of interest.[572] The combination of TRNOESY and STD-NMR has been particularly fruitful to characterize the interactions between RGD-motif peptides and various integrins, and to design peptidomimetics.[561,562,568–570,715] These were executed on live cells, but a number of studies have also used native membranes extracted from ~100.10^{6} cells. Extracted membranes permit to acquire spectra of higher quality through extended acquisition times.[716–718] This strategy has also been used for studying GPCRs in native membranes.[719–722]

It is important to note that transient, unspecific interactions between small ligands and cellular entities are very common and expected. For example, weak millimolar affinities have been detected between metabolites and *E. coli* proteins, especially for metabolites presenting aromatic rings.[723,724]

Drug candidates are often hydrophobic entities, which favor transient contacts with the countless hydrophobic patches present in and on cells. STD-NMR is also sensitive to this type of interactions, and it is advisable to acquire data also in mock-transfected cells (see Chapter 2.3.11.).

STD-NMR and TRNOESY are limited to weak-affinity interactions, in the low micromolar to millimolar range. This does not correspond to the expected affinity of any drug: FDA approved drugs have dissociation constant in the nanomolar range or below, with k_{off} between 10^{-1} and 10^{-5} s^{-1}.[725,726] However, the studies evoked above have shown that a number of interactions can be detected between native ligands and their receptor using on-cell STD-NMR effects. Hence, this capacity could be integrated in competitive assays with candidate drugs, as explicitly proposed by some of these on-cell STD studies.[398,567] These should permit to characterize drug binding to their target receptors even in cases of nano-picomolar affinities.

4.2. Drug penetrance and target engagement in live cells

Drug target engagement is a widespread issue in pharmaceutical research.[727] In the cellular context, countless potential binders can interfere, and proteins can adopt varying conformations. The upstream membrane permeability and the potential intracellular modifications are also essential parameters to evaluate. Observing drug:target interactions at atomic resolution in live cells would provide a direct readout and an efficient proof for intracellular binding and specificity. In-cell NMR spectroscopy has delivered such an information.

The most convincing examples came in the last 18 months from Banci, Luchinat and coworkers. They used transient transfection of HEK-293T cells to overexpress the carbonic anhydrase 2 (CA2) and to assess its binding to drugs supplemented in the culture medium. These interactions are revealed by chemical shift changes in the 2D ^{1}H-^{15}N NMR spectra of CA2, but also in simple one-dimensional ^{1}H spectra.[260] CA2 contains indeed four histidine residues in its enzymatic sites, whose imino groups resonate in an empty region of the ^{1}H spectrum. CA2 is expressed at 50 to 150 µM and these imino peaks are well-above the cellular background (Figure 32). In-cell one-dimensional ^{1}H spectra of CA2 provided a S/N of about 10 for acquisition times of 2-3 minutes at 150 µM in 3.10^{7} cells.[261] In a first approach, the authors exposed cell cultures expressing CA2 to varying concentrations of drug compounds, quantified the compound-bound fraction of intracellular CA2, and deduced the permeability coefficient and intracellular Kd values.[260] Interestingly, among two approved drugs treating glaucomaAAZ and MZA, AAZ diffused in cells ~15 times slower than MZA. Moreover, MZA had an apparent intracellular Kd close to its value measured *in vitro* (26 vs 14 nM). At the opposite. Another compound had an affinity 20-fold weaker in cells than *in vitro*, suggesting competing intracellular binding events. In the next report, the authors used a flow-probe bioreactor to submit cells in the NMR tube to a continuous flow of medium containing AAZ or MZA.[269] These experiments confirmed the values of membrane permeability and obtained previously. Later, the authors reported the evaluation of permeability coefficients for 9 other compounds using the first approach, i.e. in a closed tube without any flow. They noticed that this strategy can be biased by the possible release of the compound during the NMR measurement.[261] Finally, they reported another evaluation of intracellular affinities using a competition assay between ligands.[397] Here again, the use of a flow-probe made the experiments less labor-intensive, although there is an initial barrier for its implementation in a lab. This competition approach should give access to the measurement of intracellular Kd values below the nanomolar range. Altogether, Banci, Luchinat and colleagues established techniques from sample preparation to data analysis, which might be versatile enough for a number of other proteins. Applying these approaches to other proteins will confirm these promises.

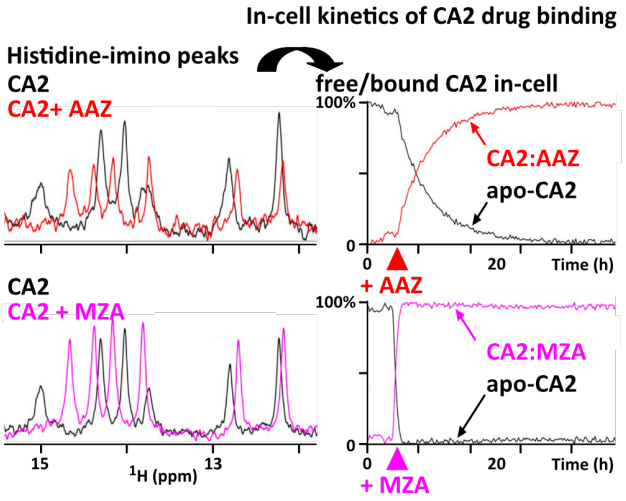


Figure 32: Time-resolved in-cell NMR monitoring of CA2 binding to approved drugs AAZ and MZA. ***Left:*** Close-up view of one dimensional in-cell ^{1}H-NMR spectra in the imino region: spectra of in-cell free CA2 (black) are overlaid with spectra of in-cell CA2:AAZ (red) and CA2:MZA (magenta) complexes.. ***Right:*** Time-resolved quantification from imino peak intensities as measured in time series of 1D ^{1}H-NMR spectra (timeframe: 7 minutes). CA2 is at 150 µM in 3.10^{7} cells, which were exposed to AAZ at 25 µM or MZA at 10 µM. Adapted from ref [269]. Copyright 2020 American Chemical Society.

The same authors have also used transfected HEK-293T cells to characterize the effect of a compound, ebselen, on the folding of ^{15}N-labeled SOD1. Using in-cell 2D ^{1}H-^{15}N NMR spectra, they could show that ebselen provoked the correct folding of mutant versions of SOD1 found in patients with familial amyotrophic lateral sclerosis (ALS).[258] They could also assess the effects of ebselen on SOD1 redox equilibrium using a flow-probe bioreactor. This set-up permitted to monitor ^{15}N-labeled SOD1 at 150 µM in 3.10^{7} cells with a time-resolution of 7 minutes.[269] The time-dependent effects of aurothioglucose (ATG), an inhibitor of thioredoxin reductase, on thioredoxin redox equilibrium have also been monitored in a flow-probe bioreactor by Shimada and colleagues (see chapter 3.3. for further explanation on this study).[326] The inhibition of protein:protein interactions can also be investigated by in-cell NMR, as shown by Shekhtman and coworkers (see chapter 3.5. for further details).[184,186,187]

DNA-binding ligands have also been tested by in-cell NMR. Trantirek and colleagues delivered pre-formed complexes, made of naphtalenophanes bound to TT mismatch containing DNA oligonucleotides, at ~10 µM in $1.3x10^{7}$ HeLa cells. One-dimensional in-cell ^{1}H-NMR spectra showed a superposition of signals from ligand-bound and

ligand-free forms for two ligands, while only the ligand-free TT-DNA was observed in presence of a third ligand.[365] This revealed the maintained capacity of the two first ligands to bind their expected targets in cells. This strategy holds true for slow binding exchange (about 1 s^{-1} or slower). These authors have shown that conformational or binding exchange in the intermediate regime (µs-ms timescale) was likely to limit the possible interpretation of in-cell NMR spectra in the case of polymorphic G-quadruplexes.[635]

Noteworthy, the NMR observation of in-cell drug engagement from the target side is favored by slow binding exchange, i.e. k_{off} rates of about 1 s^{-1} or slower. This corresponds to affinities in the range of 10-100 nM or less. Up to now, the targets have been observed preferentially at intracellular concentrations of 10-100 µM by NMR spectroscopy. This is too high for a number of reasons, and it shall be improved. However, even at these high concentrations, intracellular affinities can be measured down to ~10 nM by incubating cells with low ligand concentrations (in the range of 0.1 µM) in large extracellular volumes. Competition assays permit to measure affinities 2 order of magnitude lower, i.e. down to 100 pM. Membrane permeability, k_{on} and k_{off} rates are thus accessible. The evoked kinetics translate into extended experimental timeframe, which calls for the use of flow-probe bioreactors to maintain cells in healthy conditions during the NMR acquisition. We must mention that the question of detecting competing intracellular binding sites is not yet solved by the evoked methods. In the long-term, target-oriented NMR observation of drug binding may be also feasible in more relevant cell lines and organoids. Finally, these methods might be used to rapidly measure intracellular drug binding of disease mutants (see Chapter 4.4.).

4.3. Reporting multiple readouts: enzyme activity, metabolic status

The macromolecule-oriented NMR spectroscopy can be complemented by ligand-oriented experiments, which are tailored to monitor a reporter of an enzyme activity or the cellular metabolome. The capacities of NMR in metabolomics studies are well-known, we refer the reader to dedicated publications for further information.[399,728,729]

The best reporters of an enzyme activity are naturally its cognate substrate and products. An interesting example is meropenem, a useful antibiotic, can be hydrolyzed by the New Delhi Metallo-β-lactamase subclass 1 (NDM-1) produced by threatening antibiotic-resistant bacterial strains. It generates ^{1}H-NMR signals, whose chemical shifts change upon hydrolysis. Hence, Breeze, Hu and colleagues have recorded time series of ^{1}H-NMR spectra using suspensions of NDM-1-expressing bacteria supplemented with meropenem at 100 µM, which permitted to monitor NDM-1 activity in presence of drug-candidates (Figure 33).[541]

Fluorinated analogs of a substrate can also report for an enzyme activity, as soon as the substrate and the products show well-separated chemical shifts. Dalvit and colleagues have used a fluorinated substrate (ARN1203) of the fatty acid amide hydrolase (FAAH) overexpressed in HEK293 cells.[543] They could record ^{19}F-NMR spectra reporting for the hydrolysis ARN1203 (supplemented at 60 µM) in presence of FAAH inhibitors. In parallel, they also acquired ^{1}H-NMR spectra establishing the metabolic fingerprint of cells. The authors used only the culture supernatant in their study, because ARN1203 was able to cross freely the cell membrane. Fluorine is often present in drug compounds (~25% of approved drugs).[440] A number of ^{19}F-NMR reporters have been designed for a broad variety of enzymatic activities, and often tested in mice for MRI purposes.[730,731] ^{15}N-labeled reporters of kinases and of methionine sulfoxide reductases have also been used in cells.[277,298]

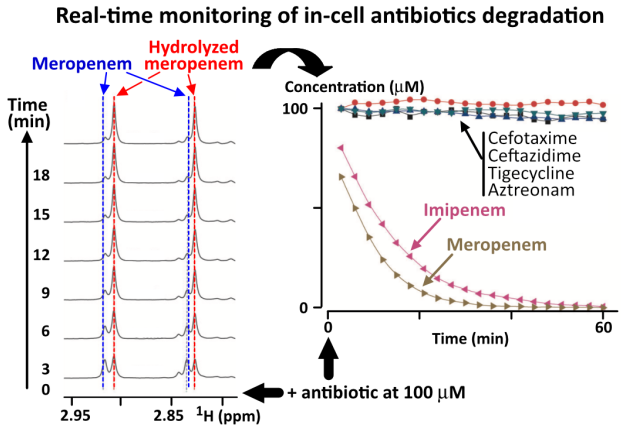


Figure 33: Example of time-resolved NMR monitoring of a cellular enzymatic activity: ***Left:*** the antibiotic meropenem is spiked in a cell suspension of *E. coli* expressing the β-lactamase NDM-1, and a time-series is recorded of one-dimensional ^{1}H-NMR spectra, where meropenem and its hydrolyzed by-product are revealed by signals at different chemical shifts; ***Right:*** the peak-intensities in every spectrum translate into molecular concentrations, which permits to monitor the reaction kinetics. Adapted from ref [541]. Copyright 2014 from Ma et al. under the terms of a Creative Commons CC BY license.

4.4. Evaluating the activity of oncogenic mutants

Characterizing, classifying and eventually interfering with oncogenic mutants is one of the goals of precision medicine.[732,733] The RAS family of GTPases is a typical target of such studies, whose mutations triggers cell proliferation and many other oncogenic signaling.[327,734] Among others, the G12, G13 and Q61 RAS mutants favor the active GTP-bound form by impairing GTP hydrolysis. The equilibrium with the inactive GDP-bound form is however regulated by RAS protein partners. This is analyzed usually by *in vitro* assays or immune-precipitation in lysates. Nishida and colleagues sought to determine this active/inactive equilibrium for RAS mutants in cells using time-resolved ^{1}H-^{13}C NMR spectroscopy. The GTP- and GDP-bound forms of RAS provoke chemical shift perturbations that can be used as spectroscopic fingerprints and permit their relative quantification. The authors delivered [^{13}C-Ile]-labeled RAS using pore-forming toxins in HeLa S3 cells, and recorded kinetics of the GTP-bound/GDP-bound equilibrium for non-farnesylated, wild-type RAS and its G12V, G13D and Q61L mutants. They showed GTP hydrolysis rates in cells than in vitro, and G12V and Q61L exhibited slower GTP-binding rates (Figure 34). Importantly, the GTP-bound/GDP-bound ratios at equilibrium were different in cells and *in vitro*, and even the order in ratios between the mutants changed in cells. These enhanced GTP-hydrolysis by RAS-G12V was surprising, because it is known to be insensitive to canonical GTPase activating proteins. Other mechanisms are thus at work, which are yet to be found.

Another interesting example was the description of SOD1 familial ALS mutants, which were not capable of adopting their native fold by themselves when expressed in HEK293 cells (see Chapter 3.1.2).[255,257] In-cell structural studies of pathogenic mutants can thus help in better understanding the deleterious mechanisms that they prompt. It would probably help to express proteins and their mutants *in situ*, in order to avoid the purification steps. These are already time-consuming in standard studies, but mutations can moreover make proteins less stable and more painful to purify. Finding set ups and labeling strategies to execute these experiments in a standardized and rapid fashion would help to establish links with clinicians.

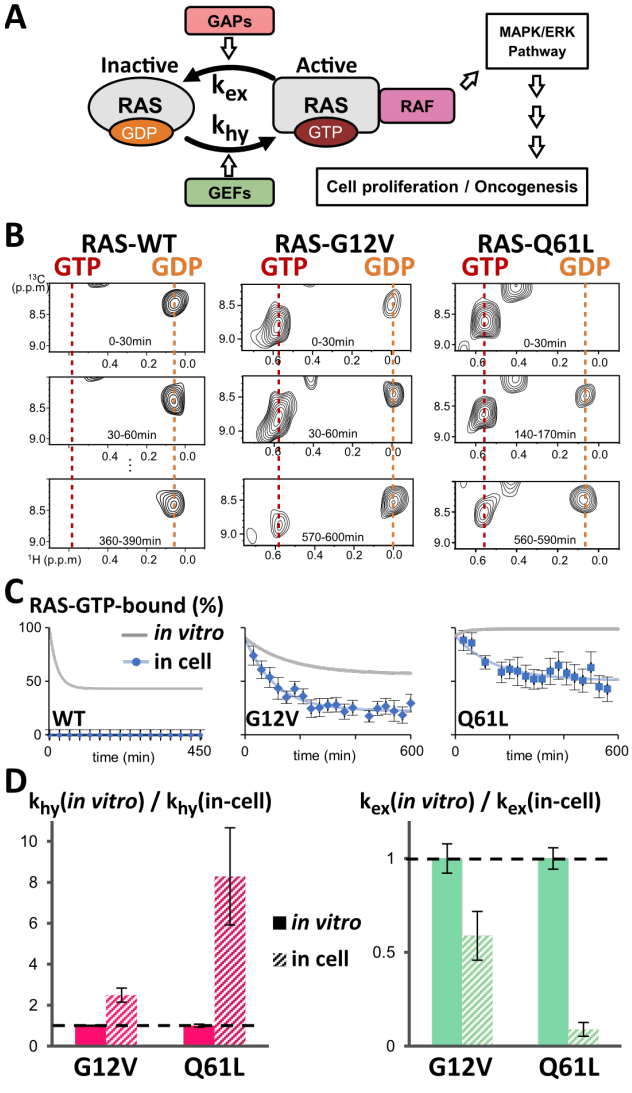


Figure 34: A) Scheme of the GTPase cycle of RAS showing its regulators GTPase-activating proteins (GAPs) and guanine nucleotide exchangers (GEFs), rates of hydrolysis (k_{hy}) and of nucleotide exchange (k_{ex}), and some of its downstream co-effectors RAF and ERK; **B)** Close-up views of two dimensional in-cell ^{1}H-^{13}C SOFAST-HMQC spectra of [u-^{2}H/^{13}CH$_3$-Ile]-labeled RAS wild-type, G12V and Q61L mutants in 3.10^7 HeLa S3 cells, showing the Ile-21 peaks sensitive to GTP- or GDP-bound states at different time points; **C)** Kinetics of GTP-bound/GDP-bound forms of RAS obtained from the peak intensities in 2D spectra shown above; **D)** Relative hydrolysis and nucleotide exchange rates measured *in vitro* and in cells for G12V and Q61L mutants of RAS. Adapted from ref [327]. Copyright 2020 from Zhao et al. under the terms of a Creative Commons CC BY license.

5- INTEGRATING IN-CELL NMR IN THE FIELD OF IN-CELL STRUCTURAL BIOLOGY

5.1. What are the main other experimental techniques for in-cell structural biology?

5.1.1. In-cell EPR

Electron Paramagnetic Resonance (EPR) spectroscopy relies on the detection of unpaired electrons, including in complex media like cells.[735,736] The resonances of unpaired electrons from "spin labels/paramagnetic tags" can deliver different types of information; we limit our list to those already reported for in-cell structural analysis.

In-cell EPR can report on local protein dynamics in the ns timescale by the analysis of the EPR line shape of nitroxide derivates, which are attached most often via a single cysteine side chain.[737] This line shape depends on the correlation time of the paramagnetic moiety. This type of experiment is performed at room temperature in a few minutes on ~10 µL samples at micromolar concentrations, using widespread continuous-wave (CW) X-band (9.8 GHz) EPR spectrometers. CW-EPR has been marginally used for in-cell studies though:[274,490,738,739] i) the obtained information is important but not very rich, ii) the spin-labels flexibility at the end of a side chain decorrelates partially its correlation time from that of the local protein dynamics, iii) and these hydrophobic spin-labels can easily interact intramolecularly or bind to cellular entities. EPR can still distinguish disordered and globular protein regions in cells,[490] or characterize α-synuclein binding to membranes.[740] It might also be capable of time-resolved monitoring of protein:protein interactions.[741]

The most appealing application of in-cell EPR is its capacity to characterize in-cell distance distribution between two spin-labels in the range of 2.5-6 nm using DEER/PELDOR or RIDME techniques.[735,742–745] They report for dipolar interactions, i.e. the sum of all the inter-label distances to the power of -3 (Σd_{ij}^{-3}). The corresponding measurements are usually executed i) during ~12 hours, ii) at cryogenic temperatures (from 10 K to 60 K, a recent trityl spin-label permitted DEER acquisition at 150 K[746]), iii) using Q-band (34 GHz) or even the less common W-band (94 GHz) spectrometers, iv) on cellular samples of ~50 µL (10-20 million mammalian cells are enough) containing low- to sub-micromolar concentrations of doubly spin-labeled proteins/nucleic acids, preferably in deuterated buffers.[391–394,742,747–749] Hence, the measured distance distribution reports for the ensemble of frozen conformations, independently of the size and location of the studied objects. The distance extraction resolves an ill-posed problem from noisy data, which can bias the distribution shape.[743,745,750–752] The global accuracy depends also on the size and rigidity of the two paramagnetic tags.[749,752–754]

In-cell EPR techniques rely on the covalent attachment of spin labels, which have to be stable in the reducing cellular environment. Resistant nitroxide labels have been designed both for proteins and nucleic acids over the last 10 years.[490,491,747,755–757] In-cell distance measurements can also

be performed between two trityl-,[392,746,748,758,759] or Gd^{3+}-based[176,390,391,393,394,739,749,760–762] The attachment of spin-labels has been mostly performed on purified proteins *in vitro*, which were later delivered in cells.[735] These paramagnetic-tag have also been directly bound on exposed loops of over-expressed membrane proteins in live cells.[742,758,763–765] In-cell EPR measurements were also carried out using genetically encodable lanthanide-binding tags,[766] or non-canonical amino acids via amber codon recoding.[738,739,767,768] Spin-labeled nanobodies might also be useful in the future.[769] The protein delivery techniques used for in-cell NMR structural studies have been adapted for in-cell EPR, notably oocyte injection,[490,740,746–748,759,762,770] or electroporation in mammalian cells.[390–394] Heat-shock was also reported recently for protein transduction into bacteria and yeast cells.[274] The latter technique is rapid, but it requires a robust thermal stability and yields rather heterogeneous protein concentrations in the cell population.

The main drawback comes from the nature of the spin-labels, which are rather large (1-2 nm) and hydrophobic entities, and of their large degrees of freedom on the residue side chains.[743,753] Moreover, EPR distance measurements requires two paramagnetic centers, which comes with potential protein destabilization and the necessity to test a number of labeling positions. Finally, the common freezing procedures are not extremely fast, in the time scale of one second. This can often bias the conformational sampling according to the few studies using sub-millisecond freezing.[771,772] Most of the in-cell EPR studies came in the last 5 years, and improved spin-labels and spectroscopic techniques are being developed to reach the full capacities of this approach.[744,745] In-cell EPR has already the proven capacity to provide nanometer-range distance distribution at sub-micromolar concentrations.[394]

5.1.2. In-cell FRET microscopy

Förster Resonance Energy Transfer (FRET)-based methods measure the energy transfer from a donor to an acceptor fluorophore, with a $<d^{-6}>$ distance dependence. It permits tag-to-tag distance measurements in the ~3-8 nm range.[773] Compared to in-cell EPR, in-cell FRET methods have the advantage of analyzing both dynamics (from the nano to the hour time-scales) and distances in live cells at physiological temperature. Moreover, they can also give access to subcellular localization of the studied objects in single cells, and they are not limited by the molecular size. As of now, in-cell FRET has only been used to derive qualitative distances associated to conformational changes, to the best of our knowledge.

In-cell FRET is carried out using two different modes: i) either on populations of a protein,[774–782] or ii) either in a single-molecule FRET (smFRET) mode.[783–787] The first approach permitted to study intracellular protein thermal stability,[774,775,778,779] or protein:protein interactions,[776,777,780] using mostly protein constructs integrating GFP derivatives and expressed *in situ*. The FRET evaluation on molecular ensembles can also be carried out after delivering proteins by microinjection.[788,789] Folding stability studies have used picosecond thermal jumps to establish the thermal stability of model proteins at ~10 μM concentrations.[774] The donor-acceptor FRET can be measured every 15 ms, and the typical unfolding kinetics last a few seconds. It does not necessarily translate directly into in-cell stability: destabilized proteins establish also rapid interactions with the chaperone Hsc70, which is present at 6 μM in mammalian cells, and eventually with the heat-inducible Hsp70.[777,780] In-cell smFRET has been used to investigate i) disordered proteins compaction, using dye-conjugated proteins microinjected in cells,[784,786] and ii) protein conformational changes upon cellular stimulation, using genetically encoded fluorophores.[785,787] To be observed as single entities, cytosolic proteins have be at nanomolar concentrations. This hampers the characterization of nano-micro-molar affinity interactions. Depending on the set-ups, the time-resolution ranges between 16 ps and 40 ms, and acquisitions last about 3 minutes before all molecules are photobleached. The fastest timeframes permits to extract the nanosecond interdye distance fluctuations, and the millisecond folding/unfolding dynamics in cells.[784,786] Indeed, in combination with Fluorescence Correlation Spectroscopy (FCS) and recurrence analysis, structural dynamics between ~10 ns to ~50 ms can be observed.[790,791] The full potential of single-molecule FRET is even more striking: multiple distance measurements permits good quality protein structural modeling (2.5-3 Å Cα-RMSD with respect to crystal structures using ~20-30 FRET restraints).[790,792,793] It might be achievable in cells, but it still requires the use of organic fluorophore tags, the genetically encoded ones being still much too large (20 to 35 kDa).[794] In combination with TIRF for membrane proteins,[787] FRET microscopy studies will probably deliver more in-cell structural information in the next years.

Here also, the main drawback comes from the necessity to attach reporting labels, whose sizes range from 1 to 5 nm. Their flexibility, their (photo)stability, their interactions with the protein surface and the necessity to deliver the labeled proteins in cells (for the non-gene-encoded fluorophores) are current limiting factors though.[754,790,795] As it is the case for EPR, FRET studies require a double-labeling by fluorophores, and thus a tedious screening of the useful and innocuous labeling positions. The set-ups are most often home-built and unique, even though the FRET community organized itself to demonstrate the reproducibility of its analysis.[774,790]

Another conformation-sensitive fluorescence method has been recently published: it exploits the conditional recognition of a peptide-tag by a fluorescent binder, which occurs upon a conformational opening of the studied protein.[796] To this aim, i) a chimera is designed to keep the peptide-tag non-accessible as long as the studied protein does not adopt an open conformation, and ii) a fluorescent binder is selected that binds the tag selectively but with a moderate affinity. Combined with TIRF and single-particle tracking, it permitted to characterize the conformational exchange of the active (open) and inactive (close) forms of the kinase Src at the membrane of transfected mammalian cells. The implementation of this technique appears to be rather tedious though: i) the peptide-tag has to be inserted in a position that permits its conformation-dependent exposure, without affecting the protein structure and dynamics; ii) it requires the fine tuning of the binder:tag affinity and of their expression levels, notably because the binder:tag interaction can readily bias the conformational ensemble of the studied protein towards the open state. The example of Src was achieved using a concentration of fluorescent-Src-tag only ~7-fold that of endogenous Src, which permitted to deliver intracellular infor-

mation of both functional and structural values. This is rare enough to be mentioned.

5.1.3. Mass-spectrometry

Structural information has been obtained at the proteomic scale in the last years. High-throughput analysis using mass-spectrometry (MS) have notably characterized proteome thermal stabilities in live cells or in tissues,[603,797–802] possibly coupled to protease sensitivity assays driven in cell lysates.[602,803] These methods rely on the assumption that proteins denature and then aggregate upon heating. Thermal proteome profiling is achieved on ~2-5.10^6 live cells or their lysates, split in a dozen of aliquots, each one being submitted during 3 minutes to a chosen temperature ranging from 37 to 75 ºC. Their quantitative MS analysis permitted the thermal characterization of thousands of proteins in cells. Such analysis can also be substituted by targeted western-blots.[804] Among others, these techniques permitted to distinguish disordered and globular proteomes,[602,803] to characterize the impact of PTMs on proteins stability,[44,805,806] to detect drug binding of targets and off-targets at the proteome scale,[797,798,801,802,807–810] to evaluate the impact of metabolic states or cell-cycle stages on proteins stability,[39,803,811,812] or to identify the binding and/or the impact of endogenous metabolites on protein stability (in lysates).[38,592,811,813] These techniques have a number of potential bias. First, MS analysis fails to detect or identify all the generated peptides, so that the protein of specific interest can be missed: e.g. more than half of the targets of FDA-approved drugs were not detected in a recent report.[603] False negatives can be numerous, e.g. if drug binding does not influence the thermal stability or the protease accessibility of the protein of interest.[797,814] It is important to note that the apparent thermal stability of a protein is often linked to that of its protein partners or complex subunits.[603,800,802] Weak interactions are moreover likely to not be detected, and the detected thermal stabilities report for an average of all possible cell localizations.[815] Organelles' separation might be possible before MS analysis,[32] but incubation at high temperatures does probably destabilize also cellular membranes and the whole cellular architecture. Are proteins in their physiological, relevant environment after 3 minutes at 60 ºC? MS studies have shown moreover how various metabolites levels, notably those of ATP, could be important for proteins stabilities.[38,592,811,813] Metabolites concentrations are probably rather affected also by high temperatures. The applications for in-cell drug binding are thus extremely effective, but the intrinsic meaning of the measured melting temperatures should be cautiously interpreted.

Structural information can also be obtained from solvent-accessibility foot-printing using chemical reactive species, and later characterized by MS analysis.[816,817] Among the approaches using this rationale, Hydrogen-Deuterium Exchange (HDX) is the most popular, but it has not been extensively used on live cells.[532,817,818] Different methods were developed and applied in cells,[817,819,820] among which fast photochemical oxidation of proteins (FPOP) appears to be promising.[817,821,822] This technique generates chemical reactions on the microsecond timescale, which can reveal the solvent-accessible residues of thousands of proteins from about 10 million cells,[823–826] or from 10,000 *Caenorhabditis elegans*.[827,828] While FPOP *in vitro* has been shown to be capable of characterizing epitope mapping, conformational changes and even of yielding good protein structure predictions,[821,829] it is still in a development phase in cells.[830,831]

Other MS proteomics strategies have been designed to map the binding-proteome of small molecules that cross-link either with reactive residues of proteins (cysteines, lysines) or upon photo-activation.[813,832–836] This makes the resulting covalent complexes resistant to the intermediate steps preparing MS analysis, i.e. cell lysis, proteolysis treatment, chromatographies… These approaches permit also competitive assays and fragment-based drug design in cells. MS-assisted drug binding was also evaluated from cells overexpressing a protein of interest, using native MS after cell lysis.[837,838]

Structural information and drug binding characterization might also be extracted from MS protein:protein interactomes obtained via biotinylation proximity assays,[33,839,840] even though these methods are biased towards IDPs.[841] MS-derived structural information has also been obtained recently using cross-linking agents (disuccinimidyl-suberate or -sulfoxide) in cells,[817,838,842–846] or in purified mitochondria.[847,848] The identified cross-linked peptides inform about their intra- and inter-molecular proximities in cells. These great achievements have been carried out upon 0.5 to 3 hours long incubation times in PBS at room temperature. This duration is enough to affect substantially the metabolites content and cell homeostasis.[267–269,395] Hence, this approach is probably more adapted to the characterization of high affinity complexes that are not affected by the cellular metabolic status.

Altogether, MS approaches provide important information that can contribute to the field of in-cell structural biology. MS permits the monitoring of "protein communities" at many levels.[849,850] This highly rich information does not provide direct in-cell information about conformations and binding, but rather from chemical modifications occurring in cells and later analyzed after cell lysis. The intermediate steps of sample handling and data treatment can generate important variations in the final results, and controversies between MS groups can emerge readily.[44,805,806,851] Finally, MS covers a detectable fraction of the proteome that does not necessarily contain the proteins of specific interest, even though overexpression would often fix this problem. Once more, the complementarity with other methods is evident.

5.1.4. Cryo-ET

The "resolution-revolution" of cryogenic electron microscopy (cryo-EM) shook the field of structural biology in the last years, where it became rapidly a dominant technique.[852] Its variant imaging method cryo-electron tomography (cryo-ET) gives access to the tri-dimensional reconstruction of cellular interior of plunged-frozen cells, and even permits the recognition of macromolecules *in situ*.[853–855] The latest technical developments coupling cryo-EM with light-microscopy and focused ion beam (FIB) milling have reinforced the capacities of cryo-ET to provide ultrastructure information.[854,856–859] Using the latest detectors, electron beam power and lamellar sample sculpting, it is possible to derive tomograms of ~100-200-nm-thin cell slices at ~3-4 nm resolution.[854,860,861] Cryo-ET studies reported in-cell structures at nanometer and subnanometer resolutions of amyloid aggregates,[862–864] tubulin- or actin-associated proteins,[865–867] or membrane protein complexes.[868–873] They were obtained using sub-tomogram averaging from cryo-ET. This approach

was initially thought to be less efficient than single-particle cryo-EM to solve structures at high resolution. However, cryo-ET is now capable to yield resolutions of 3-4 Å, thanks to the latest developments in data acquisition and computational treatment.[874–878] An outstanding 3.5 Å resolution has been reached recently for ribosomal particles in the simple *Mycoplasma pneumoniae* cells, which eventually revealed bound antibiotics.[878] Cryo-ET permitted the observation of ribosome-RNApolymerase complexes at 5-7 Å resolution also in *M. pneumoniae*. Although functionally essential, disordered regions of proteins remained "invisible" for cryo-ET, and their localization probability was modeled using distance restraints from in-cell cross-linking mass-spectrometry.[842]

In-cell cryo-ET structural determination requires the initial identification of particles in the tomograms, which is more achievable for large, abundant macromolecular complexes like ribosomes and proteasomes. To investigate a broader variety of objects in cells, it might be feasible to express chimera constructs integrating large protein scaffolds, or to co-express nanobodies-scaffolds chimeras:[879,880] these would help particle recognition, even though they would alter the physiological context. Developments for particle identity assignment are being carried out.[874,881] The recognition of molecules is however intrinsically problematic, because it relies on the extraction of expected shapes based on prior knowledge. Interactions can produce shapes that would be discarded. The principle of particle averaging and classification is moreover tailored to improve resolution, and less to detect minor populations and flexible regions.[882] Single-particle cryoEM can and already did help to solve structures down to ~50 kDa, but objects below 500 kDa are currently considered to be small for in-cell cryo-ET.[854] Improved algorithms may push back the lower molecular size limits in particle detection and classification from the noisy cellular lamellae. A limit of ~300 kDa is expected for a sample thickness of ~150 nm.[854,874,883] An important aspect about cryo-EM/ET is the sample vitrification. The current devices can freeze samples in 50 ms at best, which leaves time for conformational rearrangements and dissociation of macromolecular complexes.[772,882,884,885] The most flexible regions of proteins/nucleic acid and transient interactions will remain poorly detectable by in-cell cryo-ET in most cases.

5.2. The specific benefits of in-cell NMR

In-cell structural biology reports have multiplied only recently, mostly in the last 5 years. Hence, the full-capacities of the techniques evoked in Chapter 5.1. are yet to be established. It calls however for an evaluation of their complementarity. We discuss some of the unique aspects of in-cell NMR approaches below.

First, NMR spectroscopy is largely a label-free method, at the opposite of EPR or of fluorescence microscopy techniques. Practically, this avoids the tedious mutation screen to find innocuous labeling positions. More importantly, NMR delivers direct information on the studied object and not on the attached label. Second, in-cell solution NMR delivers information on live cells, non-destructively and possibly in a time-resolved manner. The non-destructive analysis is of particular interest for a number of questions, notably those related to redox states and metal chelation, but also for ligand binding and light-induced phenomena. It avoids the introduction of many possible biases through chemical modifications, lysis and the various separation steps that are necessary for MS analysis. Third, the complementarity of NMR spectroscopy with cryo-ET is straightforward to argue: NMR permits the characterization of flexible regions and of middle-size objects in the 20-50 kDa range, where cryo-ET goes blind.

Then, we shall acknowledge the immense advances in protein structure prediction,[4–6] and in large-scale compound screening methods, either cell-based,[886–889] or virtual.[890,891] It is now very hard to imagine how NMR would contribute more than punctually in these fields. NMR spectroscopy is a low-throughput, low-sensitivity technique *in vitro*, it does not change with cells. Concerning structural determination, prediction tools were trained to provide energetically favored structures of purified proteins. First, these can be destabilized or tuned by the intracellular environment and we need to accumulate experimental evidences in this regard. Second, these predictions tools will not be able soon to guess the intracellular populations of alternating protein/nucleic acid conformations, neither their dynamics or their variations in presence/absence of binding partners. Third, NMR spectroscopy can also provide residue-specific information on flexible/disordered regions of proteins, whose roles are essential in cell signaling, enzyme activity or amyloid diseases. Understanding all these structural behaviors will be necessary to build appropriate mental representations of the inner life of cells.

Then, concerning compound screening and target engagement, we shall underline once more that NMR spectroscopy can report for both small and large molecules at the same time. In-cell NMR can thus deliver rather comprehensive, longitudinal monitoring of drug-target pairs. Indeed, it permits the simultaneous observation of substrates, products, cofactors, metabolites, inhibitors, pH and redox potential, enzymes conformations and interactions, nucleic acids structure and modifications. This can be executed in a time-resolved fashion through days,[399,892] provided that cells are maintained in proper conditions either at low density in hydrogels or using a flow-probe bioreactor. Longitudinal monitoring is also possible with organoids, [893,894] plants,[895] or small aquatic animals like *C. elegans* or water fleas.[576,896–898] These objects generate magnetic susceptibility inhomogeneities though, which can substantially broaden signals of small molecules and call for the use of advanced pulse sequences.[586,899–901] Importantly, the direct observation of the target provides direct proofs of i) compound binding at the expected protein/nucleic acid site, and ii) of the active/unactive or the phospho/unphospho populations of the studied protein/nucleic acid.

5.3. The future technical challenges of in-cell NMR

Two important drawbacks remain: i) a lack of information on the subcellular origin of the detected signal; ii) a poor sensitivity. The first point calls for a complementary characterization by fluorescence-microscopy. It also argues in favor of spectroscopies, whose sensitivity is not affected by molecular size or binding exchange, e.g. EPR and ssNMR. All these techniques require advanced expertise, but microscopy platforms are wide-spread.

Sensitivity is the eternal Achille's heal of NMR spectroscopy. Magnetic hyperpolarization methods (DNP, para-

hydrogen) have been cheered in the last years for their theoretical game-changing signal enhancement of several ordered of magnitudes.[902–907] Their applications on live cells are however limited by the short lifetime of hyperpolarization at physiological temperature, i.e. about one minute or less.[907–910] DNP-ssNMR has been used to characterize proteins in cellular samples.[911] However, the current DNP-agent molecules are effective at ~100 K or less, which provokes low spectral resolution, high costs and technicity and limits the applications. Improved high-temperature (>200 K) DNP-agents would help to recover sharper linewidths.[912]

As far as in-cell solution NMR is concerned, investigations have been carried out using rather high concentrations of the observed species, i.e. at least 10 μM. This holds true for small and large molecules, isotope-labeled or not. This range of concentrations is much higher than endogenous concentrations for most proteins (Chapter 3.7.). Such cellular concentrations, one order of magnitude higher than native ones, perturb certainly functional interpretations. However, a number of biophysics phenomena remain roughly valid even at high concentrations, like structure, conformational dynamics, metal chelation, protein:protein affinities, ligand affinities, enzymatic activity inhibition… as long as the main cognate chaperones and binding partners of the studied proteins are present at appropriate concentrations. We summarize the important numbers very roughly below, for a ~25 kDa folded protein in ~3.10^7 mammalian cells. Numbers for bacteria, yeast and oocytes are much more variable depending on the transient, unspecific interactions with cellular components.

The current in-cell NMR concentrations permit to investigate interactions of nano- to milli-molar affinities using protein/nucleic acid-based detection, and micro- to millimolar affinities using ligand-based detection (Chapter 4). The first detection strategy relies mostly on 2D ^{1}H-^{15}N or ^{1}H-^{13}C spectra: these reveal chemical shift perturbations upon binding and are recorded at intracellular concentrations of ~25-50 μM. One-dimensional spectra might yield equivalent information using amino acid type specific ^{15}N-labeling at about twice lower concentrations. A number of strategies shall be combined to improve sensitivity: better isotope-labeling schemes, notably when introducing deuteration,[438] better pulse sequences adapted to the cellular viscosity,[439] paramagnetic T1 relaxation enhancement,[913–915] continuous readout using time-resolved non-uniform sampling,[897,916–919] deep-learning assisted spectral reconstruction,[920–922] denoising procedures,[269,923–925] deep-learning assisted peak picking,[926] … Using all these advanced techniques, it might be possible to record exploitable in-cell 2D NMR spectra of proteins/nucleic acids at intracellular concentrations of a few μmol/L.

High-quality in-cell structure determination is feasible at intracellular concentrations of ~100 μM for ~20 kDa proteins using the standard 3D spectra NOE-assisted approach.[227] The corresponding protocols might be eased by structure prediction tools.[927,928] Using pseudocontact shifts (PCSs) and residual dipolar couplings, decent in-cell structure models can be obtained at ~25 μM from 2D ^{1}H-^{15}N spectra.[287,288,362] Improved resolution would be probably achievable using supplementary ^{1}H-^{13}C-methyl PCS.[479,483] To be applied to a broad variety of proteins/nucleic acids, this approach will require improved lanthanide tags.[471,475] However, lanthanide-cage conjugation is currently operated on purified material, which must be later delivered in cells. Intracellular conjugation using non-natural amino acids would make PCS/RDC measurements also feasible with proteins expressed *in situ*.

Large conformational rearrangements can also be reported by chemical shift perturbations of a small subset of isotope-labeled residues. This corresponds to the acquisition of "simple" 2D ^{1}H-^{15}N or ^{1}H-^{13}C spectra. Large rearrangement can even be characterized by a single ^{19}F-amino acid. ^{19}F-NMR has appealing aspects because it can generate background-free one-dimensional spectra, where structural changes are exquisitely reported by chemical shift perturbations. However, the large ^{19}F CSA has largely affected in-cell ^{19}F-NMR sensitivity and linewidth until now (Chapters 2.3.4. and 2.3.7.). Advanced isotope-labeling schemes, introducing ^{13}C-^{19}F pairs, may be fruitful for in-cell NMR studies in the future.[118,119,137,139,410] Genetic code expansion may also be useful to introduce single non-natural amino acids with favorable NMR-characteristics.[120,929,930]

We feel that the approaches using *in situ* production of the analyzed proteins would favor Biologists' trust. Electroporation delivery has been shown to permit functional behavior of the transduced objects, even of complex macromolecules.[387–389,931] However, co-translational processing can have important functions in a case-dependent fashion, and it often remains as a pending question. *In situ* production would moreover avoid the time-consuming purification steps and speed-up the acquisition of the first in-cell spectral fingerprints. It would indeed certainly help to obtain fast, rough structural diagnostics about the effects of bacterial resistant mutations, or patient mutations, or drug-target engagement. Uniform ^{15}N-labeling provides rich structural information, but might be limited in terms of sensitivity in the long term. Introducing partially deuterated isotope-labeled amino acids would probably help to improve in-cell NMR sensitivity.[111,438] While *in situ* production can accommodate a number of isotope-labeling schemes in *E. coli*, there are a number of labeling possibilities to explore in insect and mammalian cells. These might also permit to use ^{13}C-direct detection, which appears to be necessary to characterize disordered proteins at physiological temperatures,[178,436] together with and ^{1}Hα-^{13}Cα correlations.[932–934] Low-gamma nuclei detection may actually be adapted to the slower tumbling times observed in cells, notably in absence of uniform deuteration. It may compete with ^{1}H-detection at GHz fields,[131,270,922,935–939] but this remains to be tested to a large extent.

In any case, in-cell NMR studies shall be executed using flow-probe bioreactors and appropriate hydrogel matrices, in order to maintain cells in healthy conditions through extended periods of times (Chapter 2.2.11.). This will also ensure reproducibility and longer acquisition times.

6. CONCLUSION

The modern in-cell structural biology by NMR was initiated about 25 years ago. It struggled for some years with poor resolution and sensitivity in most cases, notably by carrying out studies in the initial *E. coli* and frog oocytes systems. The basic mechanisms underlying these crippling intracellular relaxations are much better understood now; they are mostly due to restrained diffusion and the many transient interactions occurring at all times in a cellular environment

(Chapter 2.3.7. and 3.6.). These are actually less problematic in the less dense eukaryotic cells. Moreover, adapted isotope-labeling schemes, sample preparation and pulse sequences are progressively implemented. Hence, even though important efforts are still lying ahead to reach a micromolar sensitivity, evoking such a goal does not sound completely odd anymore.

In the meantime, although executed at intracellular concentrations of 10-200 μM, in-cell NMR of proteins/nucleic acids has been shown to be likely to tackle a number of fundamental questions, and recently to deliver promising proof-of-principles for biomedical characterizations. Hence, in-cell structural biology by NMR can cover an impressive range of subjects, like structure, conformational dynamics, folding, metal chelation, redox status, drug penetrance, drug target engagement, enzymatic activity, or mutation analysis. Because in-cell NMR is non-destructive, it has unique contributions for a number of these questions, like metal chelation or redox status, which are investigated after cell lysis when using other methods.

Time-resolved in-cell NMR allows also longitudinal and integrated monitoring of both small and large molecules involved in cellular phenomena, from drug penetrance, metabolic reactions, proteins' conformations or PTMs, etc… Altogether, in-cell NMR delivers information on cellular events from the nanosecond- to the days-timescales. This might also be achieved on organoids or small animals in the future. We shall underline once more that i) the in-cell NMR information is not localized at the subcellular level, and that ii) it can miss subpopulations of the molecules of interest, made non-detectable by promiscuous interactions with large cellular entities. The complementarity with other in-cell techniques is thus evident (see Chapters 5.1. and 5.2.).

In-cell structural biology by NMR depicts biochemistry and biophysics at work in cells, and it does so at the atomic scale. It is now ready to deliver knowledge at a faster pace, both at fundamental and applied levels. In our view, in-cell NMR is meant to be complementary to structure prediction tools and high-throughput methods. In-cell NMR generates experimental information at the atomic scale, who would prefer to miss it?

AUTHOR INFORMATION

Corresponding Author

* E-mail: francois-xavier.theillet@cnrs.fr; Phone : +3316908992.

ORCID

F.X. Theillet: 0000-0002-3264-210X

Author Contributions

The manuscript was written through contributions of all authors.

Biographies

Francois-Xavier Theillet obtained his Ph.D. degree from the University of Paris Pierre & Marie Curie (today Sorbonne Université), France, in 2010 working with Dr. Muriel Delepierre at the Institut Pasteur (Paris, France) on the structure and antigenicity of bacterial polysaccharides. He moved to the Leibniz-Institute for Molecular Pharmacology (FMP) in Berlin, where he carried out postdoctoral research with Pr. Philipp Selenko until 2015. There, he established methods for in-cell NMR studies and the characterization of post-translational modifications by NMR spectroscopy. Since 2016, he is a CNRS researcher at the Institute for Integrative Biology of the Cell (I2BC, Univ. Paris-Saclay/CEA/CNRS, Gif-sur-Yvette, France). His work aims at understanding proteins' regulation by multiple phosphorylation *in vitro* and in cells. He develops and applies biochemical and spectroscopic approaches for this purpose.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was supported by the CNRS and the CEA-Saclay, by the French Infrastructure for Integrated Structural Biology (https://www.structuralbiology.eu/networks/FRISBI, grant number ANR-10-INSB-05-01, Acronym FRISBI) and by the French National Research Agency (ANR; research grants ANR-14-ACHN-0015 and ANR-20-CE92-0013). The author thanks all his collaborators for their daily support, with a special mention to Philippe Selenko on behalf of his help in all fields through the years.

ABBREVIATIONS

CPP: cell-penetrating peptide; CSA, Chemical Shift Anisotropy; CryoEM, Cryo Electron Microscopy; CryoET, Cryo Electron Tomography; DNP, Dynamic-Nuclear Polarization; EPR, Electron Paramagnetic Resonance; FWHM: Full Width at Half Maximum; FRET, Förster/Fluorescence Resonance Energy Transfer; GFP, Green Fluorescent Protein; HDX, Hydrogen Deuterium Exchange; IDP/IDR, Intrinsically Disordered Proteins/Regions of proteins; MAS, Magic-Angle Spinning; MS, Mass Spectrometry; MW, Molecular Weight; NOE, Nuclear Overhauser Effect; OD, Optical Density; PFG, Pulsed Field Gradient; PFT: Pore-Forming Toxins; PRE, Paramagnetic Relaxation Enhancement; PCS, PseudoContact Shift; PTM, Post-Translational Modification; RDC, Residual Dipolar Coupling; SAR, Structure-Activity Relationship; SLO, Streptolysin O; S/N-PUT, Signal-to-Noise ratio Per Unit of Time; STD-NMR: Saturation Transfer Difference NMR; ssNMR, solid-state NMR; TIRF, Total Internal Reflection Microscopy.

REFERENCES

(1) Dyson, H. J.; Wright, P. E. Introduction: Biological Nuclear Magnetic Resonance. *Chem. Rev.* **2004**, *104*, 3517–3518.

(2) Lawson, C. L.; Kryshtafovych, A.; Adams, P. D.; Afonine, P. V.; Baker, M. L.; Barad, B. A.; Bond, P.; Burnley, T.; Cao, R.; Cheng, J.; et al. Cryo-EM Model Validation Recommendations Based on Outcomes of the 2019 EMDataResource Challenge. *Nat. Methods* **2021**, *18*, 156–164.

(3) Robertson, M. J.; Meyerowitz, J. G.; Skiniotis, G. Drug Discovery in the Era of Cryo-Electron Microscopy. *Trends Biochem. Sci.* **2021**.

(4) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589.

(5) Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; et al. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* **2021**, *373*, 871–876.

(6) Townshend, R. J. L.; Eismann, S.; Watkins, A. M.; Rangan, R.; Karelina, M.; Das, R.; Dror, R. O. Geometric Deep Learning of RNA Structure. *Science* **2021**, *373*, 1047–1051.

(7) Takeuchi, K.; Baskaran, K.; Arthanari, H. Structure Determination Using Solution NMR: Is It Worth the Effort? *J. Magn. Reson.* **2019**, *306*, 195–201.

(8) Wright, P. E.; Dyson, H. J. Intrinsically Disordered Proteins in Cellular Signalling and Regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29.

(9) Dyson, H. J.; Wright, P. E. Perspective: The Essential Role of NMR in the Discovery and Characterization of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2019**, *73*, 651–659.

(10) Piovesan, D.; Necci, M.; Escobedo, N.; Monzon, A. M.; Hatos, A.; Mičetić, I.; Quaglia, F.; Paladin, L.; Ramasamy, P.; Dosztanyi, Z.; et al. MobiDB: Intrinsically Disordered Proteins in 2021. *Nucleic Acids Res.* **2021**, *49*, D361–D367.

(11) Plitzko, J. M.; Schuler, B.; Selenko, P. Structural Biology Outside the Box-inside the Cell. *Curr. Opin. Struct. Biol.* **2017**, *46*, 110–121.

(12) Cramer, P. AlphaFold2 and the Future of Structural Biology. *Nat. Struct. Mol. Biol.* **2021**, *28*, 704–705.

(13) Theillet, F.-X.; Binolfi, A.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C.; Crowley, P. B.; Gierasch, L.; Pielak, G. J.; et al. Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem. Rev.* **2014**, *114*, 6661–6714.

(14) Dudev, T.; Lim, C. Competition among Metal Ions for Protein Binding Sites: Determinants of Metal Ion Selectivity in Proteins. *Chem. Rev.* **2014**, *114*, 538–556.

(15) Park, J. O.; Rubin, S. A.; Xu, Y.-F.; Amador-Noguez, D.; Fan, J.; Shlomi, T.; Rabinowitz, J. D. Metabolite Concentrations, Fluxes and Free Energies Imply Efficient Enzyme Usage. *Nat. Chem. Biol.* **2016**, *12*, 482–489.

(16) Lu, W.; Su, X.; Klein, M. S.; Lewis, I. A.; Fiehn, O.; Rabinowitz, J. D. Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. *Annu. Rev. Biochem.* **2017**, *86*, 277–304.

(17) Cheung, M. C.; LaCroix, R.; McKenna, B. K.; Liu, L.; Winkelman, J.; Ehrlich, D. J. Intracellular Protein and Nucleic Acid Measured in Eight Cell Types Using Deep-Ultraviolet Mass Mapping. *Cytometry A* **2013**, *83A*, 540–551.

(18) Wiśniewski, J. R.; Hein, M. Y.; Cox, J.; Mann, M. A "Proteomic Ruler" for Protein Copy Number and Concentration Estimation without Spike-in Standards. *Mol. Cell. Proteomics* **2014**, *13*, 3497–3506.

(19) van den Berg, J.; Boersma, A. J.; Poolman, B. Microorganisms Maintain Crowding Homeostasis. *Nat. Rev. Microbiol.* **2017**, *15*, 309–318.

(20) Ho, B.; Baryshnikova, A.; Brown, G. W. Unification of Protein Abundance Datasets Yields a Quantitative Saccharomyces Cerevisiae Proteome. *Cell Syst.* **2018**, *6*, 192-205.e3.

(21) Model, M. A.; Hollembeak, J. E.; Kurokawa, M. Macromolecular Crowding: A Hidden Link Between Cell Volume and Everything Else | Cell Physiol Biochem. *Cell. Physiol. Biochem.* **2020**, *55*, 25–40.

(22) Neurohr, G. E.; Amon, A. Relevance and Regulation of Cell Density. *Trends Cell Biol.* **2020**, *30*, 213–225.

(23) Itzhak, D. N.; Tyanova, S.; Cox, J.; Borner, G. H. Global, Quantitative and Dynamic Mapping of Protein Subcellular Localization. *eLife* **2016**, *5*, e16950.

(24) Löwe, M.; Kalacheva, M.; Boersma, A. J.; Kedrov, A. The More the Merrier: Effects of Macromolecular Crowding on the Structure and Dynamics of Biological Membranes. *FEBS J.* **2020**, *287*, 5039–5067.

(25) Carpy, A.; Krug, K.; Graf, S.; Koch, A.; Popic, S.; Hauf, S.; Macek, B. Absolute Proteome and Phosphoproteome Dynamics during the Cell Cycle of Schizosaccharomyces Pombe (Fission Yeast). *Mol. Cell. Proteomics* **2014**, *13*, 1925–1936.

(26) Sharma, K.; D'Souza, R. C. J.; Tyanova, S.; Schaab, C.; Wiśniewski, J. R.; Cox, J.; Mann, M. Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Rep.* **2014**, *8*, 1583–1594.

(27) Tsai, C.-F.; Wang, Y.-T.; Yen, H.-Y.; Tsou, C.-C.; Ku, W.-C.; Lin, P.-Y.; Chen, H.-Y.; Nesvizhskii, A. I.; Ishihama, Y.; Chen, Y.-J. Large-Scale Determination of Absolute Phosphorylation Stoichiometries in Human Cells by Motif-Targeting Quantitative Proteomics. *Nat. Commun.* **2015**, *6*, 6622.

(28) Prus, G.; Hoegl, A.; Weinert, B. T.; Choudhary, C. Analysis and Interpretation of Protein Post- Translational Modification Site Stoichiometry. *Trends Biochem. Sci.* **2019**, *44*, 943–960.

(29) Lundby, A.; Secher, A.; Lage, K.; Nordsborg, N. B.; Dmytriyev, A.; Lundby, C.; Olsen, J. V. Quantitative Maps of Protein Phosphorylation Sites across 14 Different Rat Organs and Tissues. *Nat. Commun.* **2012**, *3*, 876.

(30) Liu, Y.; Zeng, R.; Wang, R.; Weng, Y.; Wang, R.; Zou, P.; Chen, P. R. Spatiotemporally Resolved Subcellular Phosphoproteomics. *Proc. Natl. Acad. Sci.* **2021**, *118*, e2025299118.

(31) Wang, D.; Eraslan, B.; Wieland, T.; Hallström, B.; Hopf, T.; Zolg, D. P.; Zecha, J.; Asplund, A.; Li, L.; Meng, C.; et al. A Deep Proteome and Transcriptome Abundance Atlas of 29 Healthy Human Tissues. *Mol. Syst. Biol.* **2019**, *15*.

(32) Lundberg, E.; Borner, G. H. H. Spatial Proteomics: A Powerful Discovery Tool for Cell Biology. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 285–302.

(33) Go, C. D.; Knight, J. D. R.; Rajasekharan, A.; Rathod, B.; Hesketh, G. G.; Abe, K. T.; Youn, J.-Y.; Samavarchi-Tehrani, P.; Zhang, H.; Zhu, L. Y.; et al. A Proximity-Dependent Biotinylation Map of a Human Cell. *Nature* **2021**, *595*, 120–124.

(34) Rappez, L.; Stadler, M.; Triana, S.; Gathungu, R. M.; Ovchinnikova, K.; Phapale, P.; Heikenwalder, M.; Alexandrov, T. SpaceM Reveals Metabolic States of Single Cells. *Nat. Methods* **2021**, *18*, 799–805.

(35) Shibata, D.; Kajimoto, S.; Nakabayashi, T. Label-Free Tracking of Intracellular Molecular Crowding with Cell-Cycle Progression Using Raman Microscopy. *Chem. Phys. Lett.* **2021**, *779*, 138843.

(36) Rivas, G.; Minton, A. P. Macromolecular Crowding In Vitro , In Vivo , and In Between. *Trends Biochem. Sci.* **2016**, *41*, 970–981.

(37) Patel, A.; Malinovska, L.; Saha, S.; Wang, J.; Alberti, S.; Krishnan, Y.; Hyman, A. A. ATP as a Biological Hydrotrope. *Science* **2017**, *356*, 753–756.

(38) Piazza, I.; Kochanowski, K.; Cappelletti, V.; Fuhrer, T.; Noor, E.; Sauer, U.; Picotti, P. A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell* **2018**, *172*, 358-372.e23.

(39) Dai, L.; Zhao, T.; Bisteau, X.; Sun, W.; Prabhu, N.; Lim, Y. T.; Sobota, R. M.; Kaldis, P.; Nordlund, P. Modulation of Protein-Interaction States through the Cell Cycle. *Cell* **2018**, *173*, 1481-1494.e13.

(40) Owen, M. C.; Gnutt, D.; Gao, M.; Wärmländer, S. K. T. S.; Jarvet, J.; Gräslund, A.; Winter, R.; Ebbinghaus, S.; Strodel, B. Effects of *in Vivo* Conditions on Amyloid Aggregation. *Chem. Soc. Rev.* **2019**, *48*, 3946–3996.

(41) Copley, S. D. The Physical Basis and Practical Consequences of Biological Promiscuity. *Phys. Biol.* **2020**, *17*, 051001.

(42) Reshetniak, S.; Ußling, J.; Perego, E.; Rammner, B.; Schikorski, T.; Fornasiero, E. F.; Truckenbrodt, S.; Köster, S.; Rizzoli, S. O. A Comparative Analysis of the Mobility of 45 Proteins in the Synaptic Bouton. *EMBO J.* **2020**, *39*.

(43) Stadmiller, S. S.; Aguilar, J. S.; Parnham, S.; Pielak, G. J. Protein–Peptide Binding Energetics under Crowded Conditions. *J. Phys. Chem. B* **2020**, *124*, 9297–9309.

(44) Potel, C. M.; Kurzawa, N.; Becher, I.; Typas, A.; Mateus, A.; Savitski, M. M. Impact of Phosphorylation on Thermal Stability of Proteins. *Nat. Methods* **2021**, *18*, 757–759.

(45) Bah, A.; Vernon, R. M.; Siddiqui, Z.; Krzeminski, M.; Muhandiram, R.; Zhao, C.; Sonenberg, N.; Kay, L. E.; Forman-Kay, J. D. Folding of an Intrinsically Disordered Protein by Phosphorylation as a Regulatory Switch. *Nature* **2015**, *519*, 106–109.

(46) Milles, S.; Jensen, M. R.; Lazert, C.; Guseva, S.; Ivashchenko, S.; Communie, G.; Maurin, D.; Gerlier, D.; Ruigrok, R. W. H.; Blackledge, M. An Ultraweak Interaction in the Intrinsically Disordered Replication Machinery Is Essential for Measles Virus Function. *Sci. Adv.* **2018**, *4*, eaat7778.

(47) Milles, S.; Salvi, N.; Blackledge, M.; Jensen, M. R. Characterization of Intrinsically Disordered Proteins and Their Dynamic

Complexes: From in Vitro to Cell-like Environments. *Prog. Nucl. Magn. Reson. Spectrosc.* **2018**, *109*, 79–100.

(48) Jiang, Y.; Rossi, P.; Kalodimos, C. G. Structural Basis for Client Recognition and Activity of Hsp40 Chaperones. *Science* **2019**, *365*, 1313–1319.

(49) Sekhar, A.; Kay, L. E. An NMR View of Protein Dynamics in Health and Disease. *Annu. Rev. Biophys.* **2019**, *48*, 297–319.

(50) Schanda, P. Relaxing with Liquids and Solids – A Perspective on Biomolecular Dynamics. *J. Magn. Reson.* **2019**, *306*, 180–186.

(51) Xie, T.; Saleh, T.; Rossi, P.; Kalodimos, C. G. Conformational States Dynamically Populated by a Kinase Determine Its Function. *Science* **2020**, *62*, eabc2754-18.

(52) Alderson, T. R.; Kay, L. E. NMR Spectroscopy Captures the Essential Role of Dynamics in Regulating Biomolecular Function. *Cell* **2021**, *184*, 577–595.

(53) Freedberg, D. I.; Selenko, P. Live Cell NMR. *Annu. Rev. Biophys.* **2014**, *43*, 171–192.

(54) Hänsel, R.; Luh, L. M.; Corbeski, I.; Trantirek, L.; Dötsch, V. In-Cell NMR and EPR Spectroscopy of Biomacromolecules. *Angew. Chem.* **2014**, *53*, 10300–10314.

(55) Luchinat, E.; Banci, L. In-Cell NMR in Human Cells: Direct Protein Expression Allows Structural Studies of Protein Folding and Maturation. *Acc. Chem. Res.* **2018**, *51*, 1550–1557.

(56) Siegal, G.; Selenko, P. Cells, Drugs and NMR. *J. Magn. Reson.* **2019**, *306*, 202–212.

(57) Nishida, N.; Ito, Y.; Shimada, I. In Situ Structural Biology Using In-Cell NMR. *BBA - Gen. Subj.* **2019**, 1–10.

(58) Shulman, R. G. NMR Spectroscopy of Living Cells. *Sci. Am.* **1983**, *248*, 86–93.

(59) Szwergold, B. S. NMR Spectroscopy of Cells. *Annu. Rev. Physiol.* **1992**, *54*, 775–798.

(60) Serber, Z.; Dötsch, V. In-Cell NMR Spectroscopy. *Biochemistry* **2001**, *40*, 14317–14323.

(61) Lippens, G.; Cahoreau, E.; Millard, P.; Charlier, C.; Lopez, J.; Hanoulle, X.; Portais, J. C. In-Cell NMR: From Metabolites to Macromolecules. *Analyst* **2018**, *143*, 620–629.

(62) Theillet; Luchinat. In-Cell NMR: Why and How? A Long Lasting Companionship between Cells and Spins. *Prog. Nucl. Magn. Reson. Spectrosc.* **under review**.

(63) Yu, I.; Mori, T.; Ando, T.; Harada, R.; Jung, J.; Sugita, Y.; Feig, M. Biomolecular Interactions Modulate Macromolecular Structure and Dynamics in Atomistic Model of a Bacterial Cytoplasm. *eLife* **2016**, *5*, e19274.

(64) Romaniuk, J. A. H.; Cegelski, L. Bacterial Cell Wall Composition and the Influence of Antibiotics by Cell-Wall and Whole-Cell NMR. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2015**, *370*.

(65) Romaniuk, J. A. H.; Cegelski, L. Peptidoglycan and Teichoic Acid Levels and Alterations in Staphylococcus Aureus by Cell-Wall and Whole-Cell Nuclear Magnetic Resonance. *Biochemistry* **2018**, *57*, 3966–3975.

(66) Medeiros-Silva, J.; Jekhmane, S.; Paioni, A. L.; Gawarecka, K.; Baldus, M.; Swiezewska, E.; Breukink, E.; Weingarth, M. High-Resolution NMR Studies of Antibiotics in Cellular Membranes. *Nat. Commun.* **2018**, *9*, 3963–10.

(67) Kaplan, M.; Cukkemane, A.; van Zundert, G. C. P.; Narasimhan, S.; Daniëls, M.; Mance, D.; Waksman, G.; Bonvin, A. M. J. J.; Fronzes, R.; Folkers, G. E.; et al. Probing a Cell-Embedded Megadalton Protein Complex by DNP-Supported Solid-State NMR. *Nat. Methods* **2015**, *12*, 649–652.

(68) Shahid, S. A.; Nagaraj, M.; Chauhan, N.; Franks, T. W.; Bardiaux, B.; Habeck, M.; Orwick-Rydmark, M.; Linke, D.; van Rossum, B.-J. Solid-State NMR Study of the YadA Membrane-Anchor Domain in the Bacterial Outer Membrane. *Angew. Chem. Int. Ed Engl.* **2015**, *54*, 12602–12606.

(69) Kaplan, M.; Narasimhan, S.; de Heus, C.; Mance, D.; van Doorn, S.; Houben, K.; Popov-Čeleketić, D.; Damman, R.; Katrukha, E. A.; Jain, P.; et al. EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. *Cell* **2016**, *167*, 1241-1251.e11.

(70) Pinto, C.; Mance, D.; Julien, M.; Daniels, M.; Weingarth, M.; Baldus, M. Studying Assembly of the BAM Complex in Native Membranes by Cellular Solid-State NMR Spectroscopy. *J. Struct. Biol.* **2019**, *206*, 1–11.

(71) Narasimhan, S.; Scherpe, S.; Lucini Paioni, A.; van der Zwan, J.; Folkers, G. E.; Ovaa, H.; Baldus, M. DNP Supported Solid-State NMR of Proteins inside Mammalian Cells. *Angew. Chem.* **2019**, 1–7.

(72) Narasimhan, S.; Pinto, C.; Lucini Paioni, A.; van der Zwan, J.; Folkers, G. E.; Baldus, M. Characterizing Proteins in a Native Bacterial Environment Using Solid-State NMR Spectroscopy. *Nat. Protoc.* **2021**, *16*, 893–918.

(73) Biedenbänder, T.; Aladin, V.; Saeidpour, S.; Corzilius, B. Dynamic Nuclear Polarization for Sensitivity Enhancement in Biomolecular Solid-State NMR. *Chem. Rev.* **2022**, acs.chemrev.1c00776.

(74) Woodman, T. J. A Beginner's Guide to Nuclear Magnetic Resonance: From Atomic Spies to Complex 3D Structures at the Heart of Structural Biology. *The Biochemist* **2019**, *41*, 52–55.

(75) Marion, D. An Introduction to Biological NMR Spectroscopy. *Mol. Cell. Proteomics* **2013**, *12*, 3006–3025.

(76) Case, D. A. Chemical Shifts in Biomolecules. *Curr. Opin. Struct. Biol.* **2013**, *23*, 172–176.

(77) Berjanskii, M. V.; Wishart, D. S. Unraveling the Meaning of Chemical Shifts in Protein NMR. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2017**, *1865*, 1564–1576.

(78) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; et al. BioMagResBank. *Nucleic Acids Res.* **2007**, *36*, D402–D408.

(79) Mittermaier, A. K.; Kay, L. E. Observing Biological Dynamics at Atomic Resolution Using NMR. *Trends Biochem. Sci.* **2009**, *34*, 601–611.

(80) Carnevale, D.; Grosjean, B.; Bodenhausen, G. Dipolar Couplings in Solid Polypeptides Probed by 14N NMR Spectroscopy. *Commun. Chem.* **2018**, *1*, 73.

(81) Carnevale, D.; Hollenstein, M.; Bodenhausen, G. Self-Assembly of DNA and RNA Building Blocks Explored by Nitrogen-14 NMR Crystallography: Structure and Dynamics. *ChemPhysChem* **2020**, *21*, 1044–1051.

(82) Penzel, S.; Oss, A.; Org, M.-L.; Samoson, A.; Böckmann, A.; Ernst, M.; Meier, B. H. Spinning Faster: Protein NMR at MAS Frequencies up to 126 KHz. *J. Biomol. NMR* **2019**, *73*, 19–29.

(83) Boisbouvier, J. Advanced Isotopic Labeling for the NMR Investigation of Challenging Proteins and Nucleic Acids. *J. Biomol. NMR* **2018**, *71*, 115–117.

(84) Kainosho, M.; Miyanoiri, Y.; Terauchi, T.; Takeda, M. Perspective: Next Generation Isotope-Aided Methods for Protein NMR Spectroscopy. *J. Biomol. NMR* **2018**, *71*, 119–127.

(85) Schütz, S.; Sprangers, R. Methyl TROSY Spectroscopy: A Versatile NMR Approach to Study Challenging Biological Systems. *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *116*, 56–84.

(86) Kleckner, I. R.; Foster, M. P. An Introduction to NMR-Based Approaches for Measuring Protein Dynamics. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2011**, *1814*, 942–968.

(87) Anthis, N. J.; Clore, G. M. Visualizing Transient Dark States by NMR Spectroscopy. *Q. Rev. Biophys.* **2015**, *48*, 35–116.

(88) Kay, L. E. New Views of Functionally Dynamic Proteins by Solution NMR Spectroscopy. *J. Mol. Biol.* **2016**, *428*, 323–331.

(89) Alderson, T. R.; Kay, L. E. Unveiling Invisible Protein States with NMR Spectroscopy. *Curr. Opin. Struct. Biol.* **2020**, *60*, 39–49.

(90) Nishida, N.; Ito, Y.; Shimada, I. In Situ Structural Biology Using In-Cell NMR. *Biochim. Biophys. Acta Gen. Subj.* **2020**, *1864*, 129364.

(91) Hoopes, J. T.; Elberson, M. A.; Preston, R. J.; Reddy, P. T.; Kelman, Z. Protein Labeling in Escherichia Coli with 2H, 13C, and 15N. *Methods Enzymol.* **2015**, *565*, 27–44.

(92) Klopp, J.; Winterhalter, A.; Gébleux, R.; Scherer-Becker, D.; Ostermeier, C.; Gossert, A. D. Cost-Effective Large-Scale

Expression of Proteins for NMR Studies. *J. Biomol. NMR* **2018**, *71*, 247–262.

(93) Azatian, S. B.; Kaur, N.; Latham, M. P. Increasing the Buffering Capacity of Minimal Media Leads to Higher Protein Yield. *J. Biomol. NMR* **2019**, *73*, 11–17.

(94) Cai, M.; Huang, Y.; Yang, R.; Craigie, R.; Clore, G. M. A Simple and Robust Protocol for High-Yield Expression of Perdeuterated Proteins in Escherichia Coli Grown in Shaker Flasks. *J. Biomol. NMR* **2016**, *66*, 85–91.

(95) O'Brien, E. S.; Lin, D. W.; Fuglestad, B.; Stetz, M. A.; Gosse, T.; Tommos, C.; Wand, A. J. Improving Yields of Deuterated, Methyl Labeled Protein by Growing in H2O. *J. Biomol. NMR* **2018**, *71*, 263–273.

(96) Miyazawa-Onami, M.; Takeuchi, K.; Takano, T.; Sugiki, T.; Shimada, I.; Takahashi, H. Perdeuteration and Methyl-Selective 1H, 13C-Labeling by Using a Kluyveromyces Lactis Expression System. *J. Biomol. NMR* **2013**, *57*, 297–304.

(97) Fan, Y.; Emami, S.; Munro, R.; Ladizhansky, V.; Brown, L. S. Isotope Labeling of Eukaryotic Membrane Proteins in Yeast for Solid-State NMR. In *Methods in Enzymology*; Elsevier, 2015; Vol. 565, pp 193–212.

(98) Meola, A.; Deville, C.; Jeffers, S. A.; Guardado-Calvo, P.; Vasiliauskaite, I.; Sizun, C.; Girard-Blanc, C.; Malosse, C.; van Heijenoort, C.; Chamot-Rooke, J.; et al. Robust and Low Cost Uniform 15N-Labeling of Proteins Expressed in Drosophila S2 Cells and Spodoptera Frugiperda Sf9 Cells for NMR Applications. *J. Struct. Biol.* **2014**, *188*, 71–78.

(99) Sitarska, A.; Skora, L.; Klopp, J.; Roest, S.; Fernández, C.; Shrestha, B.; Gossert, A. D. Affordable Uniform Isotope Labeling with 2H, 13C and 15N in Insect Cells. *J. Biomol. NMR* **2015**, *62*, 191–197.

(100) Opitz, C.; Isogai, S.; Grzesiek, S. An Economic Approach to Efficient Isotope Labeling in Insect Cells Using Homemade 15N-, 13C- and 2H-Labeled Yeast Extracts. *J. Biomol. NMR* **2015**, *62*, 373–385.

(101) Franke, B.; Opitz, C.; Isogai, S.; Grahl, A.; Delgado, L.; Gossert, A. D.; Grzesiek, S. Production of Isotope-Labeled Proteins in Insect Cells for NMR. *J. Biomol. NMR* **2018**, *71*, 173–184.

(102) Sastry, M.; Bewley, C. A.; Kwong, P. D. *Effective Isotope Labeling of Proteins in a Mammalian Expression System*; Isotope labeling of biomolecules labeling Methods; Elsevier Inc., 2015; Vol. 565, p 307.

(103) 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.

(104) Kerfah, R.; Plevin, M. J.; Sounier, R.; Gans, P.; Boisbouvier, J. Methyl-Specific Isotopic Labeling: A Molecular Tool Box for Solution NMR Studies of Large Proteins. *Curr. Opin. Struct. Biol.* **2015**, *32*, 113–122.

(105) Rossi, P.; Monneau, Y. R.; Xia, Y.; Ishida, Y.; Kalodimos, C. G. *Toolkit for NMR Studies of Methyl-Labeled Proteins*; Isotope labeling of biomolecules labeling Methods; Elsevier Inc., 2019; p 36.

(106) Suzuki, R.; Sakakura, M.; Mori, M.; Fujii, M.; Akashi, S.; Takahashi, H. Methyl-Selective Isotope Labeling Using α-Ketoisovalerate for the Yeast Pichia Pastoris Recombinant Protein Expression System. *J. Biomol. NMR* **2018**, *71*, 213–223.

(107) Ali, R.; Clark, L. D.; Zahm, J. A.; Lemoff, A.; Ramesh, K.; Rosenbaum, D. M.; Rosen, M. K. Improved Strategy for Isoleucine 1H/13C Methyl Labeling in Pichia Pastoris. *J. Biomol. NMR* **2019**, *73*, 687–697.

(108) Zhang, M. Recent Developments of Methyl-Labeling Strategies in Pichia Pastoris for NMR Spectroscopy. *Protein Expr. Purif.* **2020**, *166*, 105521.

(109) Kofuku, Y.; Ueda, T.; Okude, J.; Shiraishi, Y.; Kondo, K.; Mizumura, T.; Suzuki, S.; Shimada, I. Functional Dynamics of Deuterated β 2-Adrenergic Receptor in Lipid Bilayers Revealed by NMR Spectroscopy. *Angew. Chem.* **2014**, *53*, 13376–13379.

(110) Kofuku, Y.; Yokomizo, T.; Imai, S.; Shiraishi, Y.; Natsume, M.; Itoh, H.; Inoue, M.; Nakata, K.; Igarashi, S.; Yamaguchi, H.; et al. Deuteration and Selective Labeling of Alanine Methyl Groups of B2-Adrenergic Receptor Expressed in a Baculovirus-Insect Cell Expression System. *J. Biomol. NMR* **2018**, *71*, 185–192.

(111) Imai, S.; Yokomizo, T.; Kofuku, Y.; Shiraishi, Y.; Ueda, T.; Shimada, I. Structural Equilibrium Underlying Ligand-Dependent Activation of B2-Adrenoreceptor. *Nat. Chem. Biol.* **2020**, *16*, 430–439.

(112) Miyanoiri, Y.; Ishida, Y.; Takeda, M.; Terauchi, T.; Inouye, M.; Kainosho, M. Highly Efficient Residue-Selective Labeling with Isotope-Labeled Ile, Leu, and Val Using a New Auxotrophic E. Coli Strain. *J. Biomol. NMR* **2016**, *65*, 109–119.

(113) Lacabanne, D. Selective Labeling and Unlabeling Strategies in Protein Solid-State NMR Spectroscopy. *J. Biomol. NMR* **2017**, *0*, 0–0.

(114) Weininger, U. Site-Selective 13C Labeling of Histidine and Tryptophan Using Ribose. *J. Biomol. NMR* **2017**, *69*, 23–30.

(115) Zhang, Y.; Wei, H.; Xie, D.; Calambur, D.; Douglas, A.; Gao, M.; Marsilio, F.; Metzler, W. J.; Szapiel, N.; Zhang, P.; et al. An Improved Protocol for Amino Acid Type-Selective Isotope Labeling in Insect Cells. *J. Biomol. NMR* **2017**, *68*, 237–247.

(116) Crowley, P. B.; Kyne, C.; Monteith, W. B. Simple and Inexpensive Incorporation of 19F-Tryptophan for Protein NMR Spectroscopy. *Chem. Commun.* **2012**, *48*, 10681–10683.

(117) Sharaf, N. G.; Gronenborn, A. M. *19F-Modified Proteins and 19F-Containing Ligands as Tools in Solution NMR Studies of Protein Interactions*; Isotope labeling of biomolecules labeling Methods; Elsevier Inc., 2015; Vol. 565, p 95.

(118) Boeszoermenyi, A.; Chhabra, S.; Dubey, A.; Radeva, D. L.; Burdzhiev, N. T.; Chanev, C. D.; Petrov, O. I.; Gelev, V. M.; Zhang, M.; Anklin, C.; et al. Aromatic 19F-13C TROSY: A Background-Free Approach to Probe Biomolecular Structure, Function, and Dynamics. *Nat. Methods* **2019**, 1–12.

(119) Maleckis, A.; Herath, I. D.; Otting, G. Synthesis of ^{13}C/^{19}F/^{2}H Labeled Indoles for Use as Tryptophan Precursors for Protein NMR Spectroscopy. *Org. Biomol. Chem.* **2021**, *19*, 5133–5147.

(120) Loh, C. T.; Adams, L. A.; Graham, B.; Otting, G. Genetically Encoded Amino Acids with Tert-Butyl and Trimethylsilyl Groups for Site-Selective Studies of Proteins by NMR Spectroscopy. *J. Biomol. NMR* **2018**, *71*, 287–293.

(121) Yokoyama, J.; Matsuda, T.; Koshiba, S.; Tochio, N.; Kigawa, T. A Practical Method for Cell-Free Protein Synthesis to Avoid Stable Isotope Scrambling and Dilution. *Anal. Biochem.* **2011**, *411*, 223–229.

(122) Terada, T.; Yokoyama, S. Escherichia Coli Cell-Free Protein Synthesis and Isotope Labeling of Mammalian Proteins. *Methods Enzymol.* **2015**, *565*, 311–345.

(123) Urbanek, A.; Morató, A.; Allemand, F.; Delaforge, E.; Fournet, A.; Popovic, M.; Delbecq, S.; Sibille, N.; Bernadó, P. A General Strategy to Access Structural Information at Atomic Resolution in Polyglutamine Homorepeats. *Angew. Chem. Int. Ed Engl.* **2018**, *57*, 3598–3601.

(124) Hoffmann, B.; Lohr, F.; LaGuerre, A.; Bernhard, F.; Dötsch, V. Protein Labeling Strategies for Liquid-State NMR Spectroscopy Using Cell-Free Synthesis. *Prog. Nucl. Magn. Reson. Spectrosc.* **2018**, *105*, 1–22.

(125) Miyanoiri, Y.; Takeda, M.; Terauchi, T.; Kainosho, M. Recent Developments in Isotope-Aided NMR Methods for Supramolecular Protein Complexes -SAIL Aromatic TROSY. *Biochim. Biophys. Acta Gen. Subj.* **2020**, *1864*, 129439.

(126) Nelissen, F. H. T.; Tessari, M.; Wijmenga, S. S.; Heus, H. A. Stable Isotope Labeling Methods for DNA. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *96*, 89–108.

(127) Fuchs, A.-L.; Neu, A.; Sprangers, R. A General Method for Rapid and Cost-Efficient Large-Scale Production of 5′ Capped RNA. *RNA* **2016**, *22*, 1454–1466.

(128) Marchanka, A.; Kreutz, C.; Carlomagno, T. Isotope Labeling for Studying RNA by Solid-State NMR Spectroscopy. *J. Biomol. NMR* **2018**, *71*, 151–164.

(129) Feyrer, H.; Munteanu, R.; Baronti, L.; Petzold, K. One-Pot Production of RNA in High Yield and Purity Through Cleaving Tandem Transcripts. *Molecules* **2020**, *25*, 1142.

(130) Karlsson, H.; Baronti, L.; Petzold, K. A Robust and Versatile Method for Production and Purification of Large-Scale RNA Samples for Structural Biology. *RNA* **2020**, *26*, 1023–1037.

(131) Schnieders, R.; Knezic, B.; Zetzsche, H.; Sudakov, A.; Matzel, T.; Richter, C.; Hengesbach, M.; Schwalbe, H.; Fürtig, B. NMR Spectroscopy of Large Functional RNAs: From Sample Preparation to Low-Gamma Detection. *Curr. Protoc. Nucleic Acid Chem.* **2020**, *82*.

(132) Nußbaumer, F.; Juen, M. A.; Gasser, C.; Kremser, J.; Müller, T.; Tollinger, M.; Kreutz, C. Synthesis and Incorporation of 13C-Labeled DNA Building Blocks to Probe Structural Dynamics of DNA by NMR. *Nucleic Acids Res.* **2017**, *45*, 9178–9192.

(133) Becette, O.; Olenginski, L. T.; Dayie, T. K. Solid-Phase Chemical Synthesis of Stable Isotope-Labeled RNA to Aid Structure and Dynamics Studies by NMR Spectroscopy. *Molecules* **2019**, *24*, 3476.

(134) Bao, H.-L.; Xu, Y. Investigation of Higher-Order RNA G-Quadruplex Structures in Vitro and in Living Cells by 19F NMR Spectroscopy. *Nat. Protoc.* **2018**, *13*, 652–665.

(135) Li, Q.; Chen, J.; Trajkovski, M.; Zhou, Y.; Fan, C.; Lu, K.; Tang, P.; Su, X.; Plavec, J.; Xi, Z.; et al. 4′-Fluorinated RNA: Synthesis, Structure, and Applications as a Sensitive 19F NMR Probe of RNA Structure and Function. *J. Am. Chem. Soc.* **2020**, *142*, 4739–4748.

(136) Himmelstoß, M.; Erharter, K.; Renard, E.; Ennifar, E.; Kreutz, C.; Micura, R. 2′- *O* -Trifluoromethylated RNA – a Powerful Modification for RNA Chemistry and NMR Spectroscopy. *Chem. Sci.* **2020**, *11*, 11322–11330.

(137) Nußbaumer, F.; Plangger, R.; Roeck, M.; Kreutz, C. Aromatic ^{19}F– ^{13}C TROSY—[^{19}F, ^{13}C]-Pyrimidine Labeling for NMR Spectroscopy of RNA. *Angew. Chem. Int. Ed.* **2020**, *59*, 17062–17069.

(138) Scott, L. G.; Hennig, M. 19F-Site-Specific-Labeled Nucleotides for Nucleic Acid Structural Analysis by NMR. In *Methods in Enzymology*; Elsevier, 2016; Vol. 566, pp 59–87.

(139) Becette, O. B.; Zong, G.; Chen, B.; Taiwo, K. M.; Case, D. A.; Dayie, T. K. Solution NMR Readily Reveals Distinct Structural Folds and Interactions in Doubly ^{13}C- and ^{19}F-Labeled RNAs. *Sci. Adv.* **2020**, *6*, eabc6572.

(140) Abramov, G.; Velyvis, A.; Rennella, E.; Wong, L. E.; Kay, L. E. A Methyl-TROSY Approach for NMR Studies of High-Molecular-Weight DNA with Application to the Nucleosome Core Particle. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 12836–12846.

(141) Williams, S. P.; Haggie, P. M.; Brindle, K. M. 19F NMR Measurements of the Rotational Mobility of Proteins in Vivo. *Biophys. J.* **1997**, *72*, 490–498.

(142) Haggie, P. M.; Brindle, K. M. Mitochondrial Citrate Synthase Is Immobilized in Vivo. *J. Biol. Chem.* **1999**, *274*, 3941–3945.

(143) 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.

(144) Jia, B.; Jeon, C. O. High-Throughput Recombinant Protein Expression in *Escherichia Coli* : Current Status and Future Perspectives. *Open Biol.* **2016**, *6*, 160196.

(145) Rosano, G. L.; Morales, E. S.; Ceccarelli, E. A. New Tools for Recombinant Protein Production in Escherichia Coli: A 5-Year Update. *Protein Sci. Publ. Protein Soc.* **2019**, *28*, 1412–1422.

(146) Serber, Z.; Selenko, P.; Hänsel, R.; Reckel, S.; Lohr, F.; Ferrell, J. E.; Wagner, G.; Dötsch, V. Investigating Macromolecules inside Cultured and Injected Cells by In-Cell NMR Spectroscopy. *Nat. Protoc.* **2006**, *1*, 2701–2709.

(147) Serber, Z.; Ledwidge, R.; Miller, S. M.; Dötsch, V. Evaluation of Parameters Critical to Observing Proteins inside Living Escherichia Coli by In-Cell NMR Spectroscopy. *J. Am. Chem. Soc.* **2001**, *123*, 8895–8901.

(148) Serber, Z.; Straub, W.; Corsini, L.; Nomura, A. M.; Shimba, N.; Craik, C. S.; Ortiz de Montellano, P.; Dötsch, V. Methyl Groups as Probes for Proteins and Complexes in In-Cell NMR Experiments. *J. Am. Chem. Soc.* **2004**, *126*, 7119–7125.

(149) Burz, D. S.; Dutta, K.; Cowburn, D.; Shekhtman, A. In-Cell NMR for Protein-Protein Interactions (STINT-NMR). *Nat. Protoc.* **2006**, *1*, 146–152.

(150) Speer, S. L.; Guseman, A. J.; Patteson, J. B.; Ehrmann, B. M.; Pielak, G. J. Controlling and Quantifying Protein Concentration in *Escherichia Coli*. *Protein Sci.* **2019**, pro.3637.

(151) Baker, L. A.; Daniëls, M.; van der Cruijsen, E. A. W.; Folkers, G. E.; Baldus, M. Efficient Cellular Solid-State NMR of Membrane Proteins by Targeted Protein Labeling. *J. Biomol. NMR* **2015**, *62*, 199–208.

(152) Reckel, S.; Lopez, J. J.; Löhr, F.; Glaubitz, C.; Dötsch, V. In-Cell Solid-State NMR as a Tool to Study Proteins in Large Complexes. *ChemBioChem* **2012**, *13*, 534–537.

(153) Dedmon, M. M.; Patel, C. N.; Young, G. B.; Pielak, G. J. FlgM Gains Structure in Living Cells. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12681–12684.

(154) McNulty, B. C.; Young, G. B.; Pielak, G. J. Macromolecular Crowding in the Escherichia Coli Periplasm Maintains α-Synuclein Disorder. *J. Mol. Biol.* **2006**, *355*, 893–897.

(155) Sharaf, N. G.; Barnes, C. O.; Charlton, L. M.; Young, G. B.; Pielak, G. J. A Bioreactor for In-Cell Protein NMR. *J. Magn. Reson.* **2010**, *202*, 140–146.

(156) Barnes, C. O.; Monteith, W. B.; Pielak, G. J. Internal and Global Protein Motion Assessed with a Fusion Construct and In-Cell NMR Spectroscopy. *ChemBioChem* **2011**, *12*, 390–391.

(157) Fu, R.; Wang, X.; Li, C.; Santiago-Miranda, A. N.; Pielak, G. J.; Tian, F. In Situ Structural Characterization of a Recombinant Protein in Native Escherichia Coli Membranes with Solid-State Magic-Angle-Spinning NMR. *J. Am. Chem. Soc.* **2011**, *133*, 12370–12373.

(158) Monteith, W. B.; Pielak, G. J. Residue Level Quantification of Protein Stability in Living Cells. *Proc. Natl. Acad. Sci.* **2014**, *111*, 11335–11340.

(159) Monteith, W. B.; Cohen, R. D.; Smith, A. E.; Guzman-Cisneros, E.; Pielak, G. J. Quinary Structure Modulates Protein Stability in Cells. *Proc. Natl. Acad. Sci.* **2015**, *112*, 1739–1742.

(160) Cohen, R. D.; Guseman, A. J.; Pielak, G. J. Intracellular PH Modulates Quinary Structure. *Protein Sci. Publ. Protein Soc.* **2015**, *24*, 1748–1755.

(161) Sciences, N. A. of. Correction for Monteith and Pielak, Residue Level Quantification of Protein Stability in Living Cells. *Proc. Natl. Acad. Sci.* **2015**, *112*, E7031–E7031.

(162) Smith, A. E.; Zhou, L. Z.; Pielak, G. J. Hydrogen Exchange of Disordered Proteins in *Escherichia Coli*: NMR Studies of Disorder in Cells. *Protein Sci.* **2015**, *24*, 706–713.

(163) Cohen, R. D.; Pielak, G. J. Electrostatic Contributions to Protein Quinary Structure. *J. Am. Chem. Soc.* **2016**, *138*, 13139–13142.

(164) Cohen, R. D.; Pielak, G. J. Quinary Interactions with an Unfolded State Ensemble: Quinary Interactions & Unfolded State Ensemble. *Protein Sci.* **2017**, *26*, 1698–1703.

(165) Hubbard, J. A.; MacLachlan, L. K.; King, G. W.; Jones, J. J.; Fosberry, A. P. Nuclear Magnetic Resonance Spectroscopy Reveals the Functional State of the Signalling Protein CheY in Vivo in Escherichia Coli. *Mol. Microbiol.* **2003**, *49*, 1191–1200.

(166) Croke, R. L.; Sallum, C. O.; Watson, E.; Watt, E. D.; Alexandrescu, A. T. Hydrogen Exchange of Monomeric α-Synuclein Shows Unfolded Structure Persists at Physiological Temperature and Is Independent of Molecular Crowding in *Escherichia Coli*. *Protein Sci.* **2008**, *17*, 1434–1445.

(167) Sakakibara, D.; Sasaki, A.; Ikeya, T.; Hamatsu, J.; Hanashima, T.; Mishima, M.; Yoshimasu, M.; Hayashi, N.; Mikawa, T.; Wälchli, M.; et al. Protein Structure Determination in Living Cells by In-Cell NMR Spectroscopy. *Nature* **2009**, *458*, 102–105.

(168) Ikeya, T.; Sasaki, A.; Sakakibara, D.; Shigemitsu, Y.; Hamatsu, J.; Hanashima, T.; Mishima, M.; Yoshimasu, M.; Hayashi, N.; Mikawa, T.; et al. NMR Protein Structure Determination in Living E. Coli Cells Using Nonlinear Sampling. *Nat. Protoc.* **2010**, *5*, 1051–1060.

(169) Ikeya, T.; Hanashima, T.; Hosoya, S.; Shimazaki, M.; Ikeda, S.; Mishima, M.; Güntert, P.; Ito, Y. Improved In-Cell Structure Determination of Proteins at near- Physiological Concentration. *Sci. Rep.* **2016**, 1–11.

(170) Crowley, P. B.; Chow, E.; Papkovskaia, T. Protein Interactions in the Escherichia Coli Cytosol: An Impediment to In-Cell NMR Spectroscopy. *ChemBioChem* **2011**, *12*, 1043–1048.

(171) Banci, L.; Barbieri, L.; Bertini, I.; Cantini, F.; Luchinat, E. In-Cell NMR in E. Coli to Monitor Maturation Steps of HSOD1. *PLoS ONE* **2011**, *6*, e23561-8.

(172) Banci, L.; Barbieri, L.; Luchinat, E.; Secci, E. Visualization of Redox-Controlled Protein Fold in Living Cells. *Chem. Biol.* **2013**, *20*, 747–752.

(173) Wang, Q.; Zhuravleva, A.; Gierasch, L. M. Exploring Weak, Transient Protein--Protein Interactions in Crowded in Vivo Environments by in-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2011**, *50*, 9225–9236.

(174) Waudby, C. A.; Mantle, M. D.; Cabrita, L. D.; Gladden, L. F.; Dobson, C. M.; Christodoulou, J. Rapid Distinction of Intracellular and Extracellular Proteins Using NMR Diffusion Measurements. *J. Am. Chem. Soc.* **2012**, *134*, 11312–11315.

(175) Waudby, C. A.; Camilloni, C.; Fitzpatrick, A. W. P.; Cabrita, L. D.; Dobson, C. M.; Vendruscolo, M.; Christodoulou, J. In-Cell NMR Characterization of the Secondary Structure Populations of a Disordered Conformation of α-Synuclein within E. Coli Cells. *PLoS ONE* **2013**, *8*, e72286.

(176) Theillet, F.-X.; Binolfi, A.; Bekei, B.; Martorana, A.; Rose, H. M.; Stuiver, M.; Verzini, S.; Lorenz, D.; van Rossum, M.; Goldfarb, D.; et al. Structural Disorder of Monomeric α-Synuclein Persists in Mammalian Cells. *Nature* **2016**, *530*, 45–50.

(177) Binolfi, A.; Theillet, F.-X.; Selenko, P. Bacterial In-Cell NMR of Human α-Synuclein: A Disordered Monomer by Nature? *Biochem. Soc. Trans.* **2012**, *40*, 950–954.

(178) Felli, I. C.; Gonnelli, L.; Pierattelli, R. In-Cell 13C NMR Spectroscopy for the Study of Intrinsically Disordered Proteins. *Nat. Protoc.* **2014**, *9*, 2005–2016.

(179) Hough, L. E.; Dutta, K.; Sparks, S.; Temel, D. B.; Kamal, A.; Tetenbaum-Novatt, J.; Rout, M. P.; Cowburn, D. The Molecular Mechanism of Nuclear Transport Revealed by Atomic-Scale Measurements. *eLife* **2015**, *4*, e10027.

(180) Popovic, M.; Sanfelice, D.; Pastore, C.; Prischi, F.; Temussi, P. A.; Pastore, A. Selective Observation of the Disordered Import Signal of a Globular Protein by In-Cell NMR: The Example of Frataxins. *Protein Sci. Publ. Protein Soc.* **2015**, *24*, 996–1003.

(181) Song, X.; Lv, T.; Chen, J.; Wang, J.; Yao, L. Characterization of Residue Specific Protein Folding and Unfolding Dynamics in Cells. *J. Am. Chem. Soc.* **2019**, *141*, 11363–11366.

(182) 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.

(183) Burz, D. S.; Shekhtman, A. In-Cell Biochemistry Using NMR Spectroscopy. *PLoS ONE* **2008**, *3*, e2571.

(184) Xie, J.; Thapa, R.; Reverdatto, S.; Burz, D. S.; Shekhtman, A. Screening of Small Molecule Interactor Library by Using In-Cell NMR Spectroscopy (SMILI-NMR). *J. Med. Chem.* **2009**, *52*, 3516–3522.

(185) Maldonado, A. Y.; Burz, D. S.; Reverdatto, S.; Shekhtman, A. Fate of Pup inside the Mycobacterium Proteasome Studied by In-Cell NMR. *PLoS ONE* **2013**, *8*, e74576-16.

(186) Cobbert, J. D.; DeMott, C.; Majumder, S.; Smith, E. A.; Reverdatto, S.; Burz, D. S.; McDonough, K. A.; Shekhtman, A. Caught in Action: Selecting Peptide Aptamers against Intrinsically Disordered Proteins in Live Cells. *Sci. Rep.* **2015**, *5*, 9402.

(187) DeMott, C. M.; Girardin, R.; Cobbert, J.; Reverdatto, S.; Burz, D. S.; McDonough, K.; Shekhtman, A. Potent Inhibitors of Mycobacterium Tuberculosis Growth Identified by Using In-Cell NMR-Based Screening. *ACS Chem. Biol.* **2018**, *13*, 733–741.

(188) Shimba, N.; Serber, Z.; Ledwidge, R.; Miller, S. M.; Craik, C. S.; Dötsch, V. Quantitative Identification of the Protonation State of Histidines in Vitro and in Vivo †. *Biochemistry* **2003**, *42*, 9227–9234.

(189) Xu, G.; Ye, Y.; Liu, X.; Cao, S.; Wu, Q.; Cheng, K.; Liu, M.; Pielak, G. J.; Li, C. Strategies for Protein NMR in Escherichia Coli. *Biochemistry* **2014**, *53*, 1971–1981.

(190) Burz, D. S.; Breindel, L.; Shekhtman, A. Improved Sensitivity and Resolution of In-Cell NMR Spectra. In *Chemical and Synthetic Biology Approaches To Understand Cellular Functions - Part A*; Elsevier, 2019; Vol. 621, pp 305–328.

(191) Jackson, J. C.; Hammill, J. T.; Mehl, R. A. Site-Specific Incorporation of a 19F-Amino Acid into Proteins as an NMR Probe for Characterizing Protein Structure and Reactivity. *J. Am. Chem. Soc.* **2007**, *129*, 1160–1166.

(192) Ye, Y.; Liu, X.; Xu, G.; Liu, M.; Li, C. Direct Observation of Ca2+-Induced Calmodulin Conformational Transitions in Intact Xenopus Laevis Oocytes by 19F NMR Spectroscopy. *Angew. Chem.* **2015**, *127*, 5418–5420.

(193) 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.

(194) Schlesinger, A. P.; Wang, Y.; Tadeo, X.; Millet, O.; Pielak, G. J. Macromolecular Crowding Fails To Fold a Globular Protein in Cells. *J. Am. Chem. Soc.* **2011**, *133*, 8082–8085.

(195) Ye, Y.; Liu, X.; Zhang, Z.; Wu, Q.; Jiang, B.; Jiang, L.; Zhang, X.; Liu, M.; Pielak, G. J.; Li, C. 19F NMR Spectroscopy as a Probe of Cytoplasmic Viscosity and Weak Protein Interactions in Living Cells. *Chem. - Eur. J.* **2013**, *19*, 12705–12710.

(196) Speer, S. L.; Zheng, W.; Jiang, X.; Chu, I.-T.; Guseman, A. J.; Liu, M.; Pielak, G. J.; Li, C. The Intracellular Environment Affects Protein–Protein Interactions. *Proc. Natl. Acad. Sci.* **2021**, *118*, e2019918118.

(197) Ye, Y.; Wu, Q.; Zheng, W.; Jiang, B.; Pielak, G. J.; Liu, M.; Li, C. Positively Charged Tags Impede Protein Mobility in Cells as Quantified by ^{19} F NMR. *J. Phys. Chem. B* **2019**, *123*, 4527–4533.

(198) Stadmiller, S. S.; Gorensek-Benitez, A. H.; Guseman, A. J.; Pielak, G. J. Osmotic Shock Induced Protein Destabilization in Living Cells and Its Reversal by Glycine Betaine. *J. Mol. Biol.* **2017**, *429*, 1155–1161.

(199) Ye, Y.; Wu, Q.; Zheng, W.; Jiang, B.; Pielak, G. J.; Liu, M.; Li, C. Quantification of Size Effect on Protein Rotational Mobility in Cells by 19F NMR Spectroscopy. *Anal. Bioanal. Chem.* **2018**, *410*, 869–874.

(200) Smith, A. E.; Zhou, L. Z.; Gorensek, A. H.; Senske, M.; Pielak, G. J. In-Cell Thermodynamics and a New Role for Protein Surfaces. *Proc. Natl. Acad. Sci.* **2016**, *113*, 1725–1730.

(201) Kyne, C.; Crowley, P. B. Short Arginine Motifs Drive Protein Stickiness in the *Escherichia Coli* Cytoplasm. *Biochemistry* **2017**, *56*, 5026–5032.

(202) Chu, I.-T.; Speer, S. L.; Pielak, G. J. Rheostatic Control of Protein Expression Using Tuner Cells. *Biochemistry* **2020**, *59*, 733–735.

(203) Cruzeiro-Silva, C.; Albernaz, F. P.; Valente, A. P.; Almeida, F. C. L. In-Cell NMR Spectroscopy: Inhibition of Autologous Protein Expression Reduces Escherichia Coli Lysis. *Cell Biochem. Biophys.* **2006**, *44*, 497–502.

(204) Barnes, C. O.; Pielak, G. J. In-Cell Protein NMR and Protein Leakage. *Proteins Struct. Funct. Bioinforma.* **2011**, *79*, 347–351.

(205) Guseman, A. J.; Pielak, G. J. Chapter 12:Protein Stability and Weak Intracellular Interactions. In *In-cell NMR Spectroscopy*; 2019; pp 188–206.

(206) Sugiki, T.; Yamaguchi, Y.; Fujiwara, T.; Inouye, M.; Ito, Y.; Kojima, C. In-Cell NMR as a Sensitive Tool to Monitor Physiological Condition of Escherichia Coli. *Sci. Rep.* **2020**, *10*, 2466–2468.

(207) Lozano Terol, G.; Gallego-Jara, J.; Sola Martínez, R. A.; Martínez Vivancos, A.; Cánovas Díaz, M.; de Diego Puente, T. Impact of the Expression System on Recombinant Protein Production in Escherichia Coli BL21. *Front. Microbiol.* **2021**, *12*, 682001.

(208) Borkowski, O.; Ceroni, F.; Stan, G.-B.; Ellis, T. Overloaded and Stressed: Whole-Cell Considerations for Bacterial Synthetic Biology. *Curr. Opin. Microbiol.* **2016**, *33*, 123–130.

(209) Dvorak, P.; Chrast, L.; Nikel, P. I.; Fedr, R.; Soucek, K.; Sedlackova, M.; Chaloupkova, R.; de Lorenzo, V.; Prokop, Z.; Damborsky, J. Exacerbation of Substrate Toxicity by IPTG in Escherichia Coli BL21(DE3) Carrying a Synthetic Metabolic Pathway. *Microb. Cell Factories* **2015**, *14*, 201.

(210) Haddadin, F. T.; Harcum, S. W. Transcriptome Profiles for High-Cell-Density Recombinant and Wild-Type *Escherichia Coli*. *Biotechnol. Bioeng.* **2005**, *90*, 127–153.

(211) Dürrschmid, K.; Reischer, H.; Schmidt-Heck, W.; Hrebicek, T.; Guthke, R.; Rizzi, A.; Bayer, K. Monitoring of Transcriptome and Proteome Profiles to Investigate the Cellular Response of E. Coli towards Recombinant Protein Expression under Defined Chemostat Conditions. *J. Biotechnol.* **2008**, *135*, 34–44.

(212) James, J.; Yarnall, B.; Koranteng, A.; Gibson, J.; Rahman, T.; Doyle, D. A. Protein Over-Expression in Escherichia Coli Triggers Adaptation Analogous to Antimicrobial Resistance. *Microb. Cell Factories* **2021**, *20*, 13.

(213) Wang, G.; Huang, M.; Nielsen, J. Exploring the Potential of Saccharomyces Cerevisiae for Biopharmaceutical Protein Production. *Curr. Opin. Biotechnol.* **2017**, *48*, 77–84.

(214) Vieira Gomes, A.; Souza Carmo, T.; Silva Carvalho, L.; Mendonça Bahia, F.; Parachin, N. Comparison of Yeasts as Hosts for Recombinant Protein Production. *Microorganisms* **2018**, *6*, 38.

(215) Mohammadzadeh, R.; Karbalaei, M.; Soleimanpour, S.; Mosavat, A.; Rezaee, S. A.; Ghazvini, K.; Farsiani, H. Practical Methods for Expression of Recombinant Protein in the *Pichia Pastoris* System. *Curr. Protoc.* **2021**, *1*.

(216) Vidal, M.; Fields, S. The Yeast Two-Hybrid Assay: Still Finding Connections after 25 Years. *Nat. Methods* **2014**, *11*, 1203–1206.

(217) 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.

(218) Narayanaswamy, R.; Levy, M.; Tsechansky, M.; Stovall, G. M.; O'Connell, J. D.; Mirrielees, J.; Ellington, A. D.; Marcotte, E. M. Widespread Reorganization of Metabolic Enzymes into Reversible Assemblies upon Nutrient Starvation. *Proc. Natl. Acad. Sci.* **2009**, *106*, 10147–10152.

(219) Sun, S.; Gresham, D. Cellular Quiescence in Budding Yeast. *Yeast* **2021**, *38*, 12–29.

(220) Shimada, I.; Ueda, T.; Kofuku, Y.; Eddy, M. T.; Wüthrich, K. GPCR Drug Discovery: Integrating Solution NMR Data with Crystal and Cryo-EM Structures. *Nat. Rev. Drug Discov.* **2018**, 1–24.

(221) Nygaard, R.; Zou, Y.; Dror, R. O.; Mildorf, T. J.; Arlow, D. H.; Manglik, A.; Pan, A. C.; Liu, C. W.; Fung, J. J.; Bokoch, M. P.; et al. The Dynamic Process of β(2)-Adrenergic Receptor Activation. *Cell* **2013**, *152*, 532–542.

(222) Isogai, S.; Deupi, X.; Opitz, C.; Heydenreich, F. M.; Tsai, C.-J.; Brueckner, F.; Schertler, G. F. X.; Veprintsev, D. B.; Grzesiek, S. Backbone NMR Reveals Allosteric Signal Transduction Networks in the B1-Adrenergic Receptor. *Nature* **2016**, *530*, 237–241.

(223) Kleist, A. B.; Peterson, F.; Tyler, R. C.; Gustavsson, M.; Handel, T. M.; Volkman, B. F. Solution NMR Spectroscopy of GPCRs: Residue-Specific Labeling Strategies with a Focus on 13C-Methyl Methionine Labeling of the Atypical Chemokine Receptor ACKR3. *Methods Cell Biol.* **2019**, *149*, 259–288.

(224) Kanev, G. K.; de Graaf, C.; Westerman, B. A.; de Esch, I. J.; Kooistra, A. J. KLIFS: An Overhaul after the First 5 Years of Supporting Kinase Research. *Nucleic Acids Res.* **2020**, *49*, D562–D569.

(225) Errey, J. C.; Fiez-Vandal, C. Production of Membrane Proteins in Industry: The Example of GPCRs. *Protein Expr. Purif.* **2020**, *169*, 105569.

(226) Hamatsu, J.; O'Donovan, D.; Tanaka, T.; Shirai, T.; Hourai, Y.; Mikawa, T.; Ikeya, T.; Mishima, M.; Boucher, W.; Smith, B. O.; et al. 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.

(227) Tanaka, T.; Ikeya, T.; Kamoshida, H.; Suemoto, Y.; Mishima, M.; Shirakawa, M.; Güntert, P.; Ito, Y. High-Resolution Protein 3D Structure Determination in Living Eukaryotic Cells. *Angew. Chem.* **2019**, *58*, 7284–7288.

(228) McKenzie, E. A.; Abbott, W. M. Expression of Recombinant Proteins in Insect and Mammalian Cells. *Methods San Diego Calif* **2018**, *147*, 40–49.

(229) Martínez-Solís, M.; Herrero, S.; Targovnik, A. M. Engineering of the Baculovirus Expression System for Optimized Protein Production. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 113–123.

(230) Monteiro, F.; Carinhas, N.; Carrondo, M. J. T.; Bernal, V.; Alves, P. M. Toward System-Level Understanding of Baculovirus–Host Cell Interactions: From Molecular Fundamental Studies to Large-Scale Proteomics Approaches. *Front. Microbiol.* **2012**, *3*.

(231) Zhang, X.; Xu, K.; Wei, D.; Wu, W.; Yang, K.; Yuan, M. Baculovirus Infection Induces Disruption of the Nuclear Lamina. *Sci. Rep.* **2017**, *7*, 7823.

(232) Bleckmann, M.; Schürig, M.; Endres, M.; Samuels, A.; Gebauer, D.; Konisch, N.; van den Heuvel, J. Identifying Parameters to Improve the Reproducibility of Transient Gene Expression in High Five Cells. *PLOS ONE* **2019**, *14*, e0217878.

(233) Takeuchi, K.; Baskaran, K.; Arthanari, H. Structure Determination Using Solution NMR: Is It Worth the Effort? *J. Magn. Reson.* **2019**, *306*, 195–201.

(234) London, R. E.; Gregg, C. T.; Matwiyoff, N. A. Nuclear Magnetic Resonance of Rotational Mobility of Mouse Hemoglobin Labeled with (2-13C)Histidine. *Science* **1975**, *188*, 266–268.

(235) Peterson, J. L.; McConkey, E. H. Proteins of Friend Leukemia Cells. Comparison of Hemoglobin-Synthesizing and Noninduced Populations. *J. Biol. Chem.* **1976**, *251*, 555–558.

(236) Lapidot, A.; Irving, C. S.; Malik, Z. The 15N Nuclear Magnetic Resonance of Friend Leukemic Cell (Gly-15N) Hemoblogin. The Resolution of Noncovalent Bonding Interactions. *J. Am. Chem. Soc.* **1976**, *98*, 632–634.

(237) Barbieri, L.; Bertini, I.; Luchinat, E.; Secci, E.; Zhao, Y.; Banci, L.; Aricescu, A. R. Atomic-Resolution Monitoring of Protein Maturation in Live Human Cells by NMR. *Nat. Chem. Biol.* **2013**, *9*, 297–299.

(238) Barbieri, L.; Luchinat, E.; Banci, L. Characterization of Proteins by In-Cell NMR Spectroscopy in Cultured Mammalian Cells. *Nat. Protoc.* **2016**, *11*, 1101–1111.

(239) Aricescu, A. R.; Assenberg, R.; Bill, R. M.; Busso, D.; Chang, V. T.; Davis, S. J.; Dubrovsky, A.; Gustafsson, L.; Hedfalk, K.; Heinemann, U.; et al. Eukaryotic Expression: Developments for Structural Proteomics. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 1114–1124.

(240) Aricescu, A. R.; Owens, R. J. Expression of Recombinant Glycoproteins in Mammalian Cells: Towards an Integrative Approach to Structural Biology. *Curr. Opin. Struct. Biol.* **2013**, *23*, 345–356.

(241) Stepanenko, A. A.; Dmitrenko, V. V. HEK293 in Cell Biology and Cancer Research: Phenotype, Karyotype, Tumorigenicity, and Stress-Induced Genome-Phenotype Evolution. *Gene* **2015**, *569*, 182–190.

(242) Pulix, M.; Lukashchuk, V.; Smith, D. C.; Dickson, A. J. Molecular Characterization of HEK293 Cells as Emerging Versatile Cell Factories. *Curr. Opin. Biotechnol.* **2021**, *71*, 18–24.

(243) Fischer, S.; Handrick, R.; Otte, K. The Art of CHO Cell Engineering: A Comprehensive Retrospect and Future Perspectives. *Biotechnol. Adv.* **2015**, *33*, 1878–1896.

(244) Hong, J. K.; Lakshmanan, M.; Goudar, C.; Lee, D.-Y. Towards next Generation CHO Cell Line Development and Engineering by Systems Approaches. *Curr. Opin. Chem. Eng.* **2018**, *22*, 1–10.

(245) Walsh, G. Biopharmaceutical Benchmarks 2018. *Nat. Biotechnol.* **2018**, *36*, 1136–1145.

(246) Tihanyi, B.; Nyitray, L. Recent Advances in CHO Cell Line Development for Recombinant Protein Production. *Drug Discov. Today Technol.* **2021**, S1740674921000068.

(247) Nettleship, J. E.; Watson, P. J.; Rahman-Huq, N.; Fairall, L.; Posner, M. G.; Upadhyay, A.; Reddivari, Y.; Chamberlain, J. M. G.; Kolstoe, S. E.; Bagby, S.; et al. Transient Expression in HEK 293 Cells: An Alternative to E. Coli for the Production of Secreted and Intracellular Mammalian Proteins. In *Insoluble Proteins*; García-Fruitós, E., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2015; Vol. 1258, pp 209–222.

(248) Nigi, I.; Fairall, L.; Schwabe, J. W. R. Expression and Purification of Protein Complexes Suitable for Structural Studies Using Mammalian HEK 293F Cells. *Curr. Protoc. Protein Sci.* **2017**, *90*.

(249) Elegheert, J.; Behiels, E.; Bishop, B.; Scott, S.; Woolley, R. E.; Griffiths, S. C.; Byrne, E. F. X.; Chang, V. T.; Stuart, D. I.; Jones, E. Y.; et al. Lentiviral Transduction of Mammalian Cells for Fast, Scalable and High-Level Production of Soluble and Membrane Proteins. *Nat. Protoc.* **2018**, *13*, 2991–3017.

(250) Abaandou, L.; Quan, D.; Shiloach, J. Affecting HEK293 Cell Growth and Production Performance by Modifying the Expression of Specific Genes. *Cells* **2021**, *10*, 1667.

(251) Hsieh, C.-L.; Goldsmith, J. A.; Schaub, J. M.; DiVenere, A. M.; Kuo, H.-C.; Javanmardi, K.; Le, K. C.; Wrapp, D.; Lee, A. G.; Liu, Y.; et al. Structure-Based Design of Prefusion-Stabilized SARS-CoV-2 Spikes. *Science* **2020**, *369*, 1501–1505.

(252) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM Structure of the 2019-NCoV Spike in the Prefusion Conformation. *Science* **2020**, *367*, 1260–1263.

(253) Zhou, D.; Duyvesteyn, H. M. E.; Chen, C.-P.; Huang, C.-G.; Chen, T.-H.; Shih, S.-R.; Lin, Y.-C.; Cheng, C.-Y.; Cheng, S.-H.; Huang, Y.-C.; et al. Structural Basis for the Neutralization of SARS-CoV-2 by an Antibody from a Convalescent Patient. *Nat. Struct. Mol. Biol.* **2020**, *27*, 950–958.

(254) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-Specific Glycan Analysis of the SARS-CoV-2 Spike. *Science* **2020**, *91*, eabb9983-333.

(255) 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.

(256) Barbieri, L.; Luchinat, E.; Banci, L. Protein Interaction Patterns in Different Cellular Environments Are Revealed by In-Cell NMR. *Sci. Rep.* **2015**, 1–9.

(257) Luchinat, E.; Barbieri, L.; Banci, L. A Molecular Chaperone Activity of CCS Restores the Maturation of SOD1 FALS Mutants. *Sci. Rep.* **2017**, 1–8.

(258) Capper, M. J.; Wright, G. S. A.; Barbieri, L.; Luchinat, E.; Mercatelli, E.; McAlary, L.; Yerbury, J. J.; O'Neill, P. M.; Antonyuk, S. V.; Banci, L.; et al. The Cysteine-Reactive Small Molecule Ebselen Facilitates Effective SOD1 Maturation. *Nat. Commun.* **2018**, *9*, 1693–1699.

(259) Polykretis, P.; Cencetti, F.; Donati, C.; Luchinat, E.; Banci, L. Cadmium Effects on Superoxide Dismutase 1 in Human Cells Revealed by NMR. *Redox Biol.* **2019**, *21*, 101102.

(260) Luchinat, E.; Barbieri, L.; Cremonini, M.; Nocentini, A.; Supuran, C. T.; Banci, L. Drug Screening in Human Cells by NMR Spectroscopy Allows the Early Assessment of Drug Potency. *Angew. Chem.* **2020**, *59*, 6535–6539.

(261) Luchinat, E.; Barbieri, L.; Cremonini, M.; Nocentini, A.; Supuran, C. T.; Banci, L. Intracellular Binding/Unbinding Kinetics of Approved Drugs to Carbonic Anhydrase II Observed by in-Cell NMR. *ACS Chem. Biol.* **2020**, *15*, 2792–2800.

(262) Salazar, A.; Keusgen, M.; Hagen, J. Amino Acids in the Cultivation of Mammalian Cells. *Amino Acids* **2016**, *48*, 1161–1171.

(263) Lee, K.; Lee, J. H. Stable Isotope Labeling of Proteins in Mammalian Cells. *J. Korean Magn. Reson. Soc.* **2020**, *24*, 77–85.

(264) Sastry, M.; Bewley, C. A.; Kwong, P. D. *Effective Isotope Labeling of Proteins in a Mammalian Expression System*; Isotope labeling of biomolecules labeling Methods; Elsevier Inc., 2015; Vol. 565, p 307.

(265) Chen, X.; Wei, S.; Ji, Y.; Guo, X.; Yang, F. Quantitative Proteomics Using SILAC: Principles, Applications, and Developments. *PROTEOMICS* **2015**, *15*, 3175–3192.

(266) Luchinat, E.; Secci, E.; Cencetti, F.; Bruni, P. Sequential Protein Expression and Selective Labeling for In-Cell NMR in Human Cells. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2016**, *1860*, 527–533.

(267) 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 Engl.* **2013**, *52*, 1208–1211.

(268) Cerofolini, L.; Giuntini, S.; Barbieri, L.; Pennestri, M.; Codina, A.; Fragai, M.; Banci, L.; Luchinat, E.; Ravera, E. Real-Time Insights into Biological Events: In-Cell Processes and Protein-Ligand Interactions. *Biophys. J.* **2019**, *116*, 239–247.

(269) Luchinat, E.; Barbieri, L.; Campbell, T. F.; Banci, L. Real-Time Quantitative In-Cell NMR: Ligand Binding and Protein Oxidation Monitored in Human Cells Using Multivariate Curve Resolution. *Anal. Chem.* **2020**, *92*, 9997–10006.

(270) Luchinat, E.; Barbieri, L.; Cremonini, M.; Banci, L. Protein In-Cell NMR Spectroscopy at 1.2 GHz. *J. Biomol. NMR* **2021**.

(271) Bekei, B.; Rose, H. M.; Herzig, M.; Dose, A.; Schwarzer, D.; Selenko, P. In-Cell NMR in Mammalian Cells: Part 1. *Intrinsically Disord. Protein Anal.* **2012**, 43–54.

(272) Bekei, B.; Rose, H. M.; Herzig, M.; Selenko, P. In-Cell NMR in Mammalian Cells: Part 2. *Intrinsically Disord. Protein Anal.* **2012**, 55–66.

(273) Bekei, B.; Rose, H. M.; Herzig, M.; Stephanowitz, H.; Krause, E.; Selenko, P. In-Cell NMR in Mammalian Cells: Part 3. *Intrinsically Disord. Protein Anal.* **2012**, 67–83.

(274) Torricella, F.; Bonucci, A.; Polykretis, P.; Cencetti, F.; Banci, L. Rapid Protein Delivery to Living Cells for Biomolecular Investigation. *Biochem. Biophys. Res. Commun.* **2021**, *570*, 82–88.

(275) Stewart, M. P.; Langer, R.; Jensen, K. F. Intracellular Delivery by Membrane Disruption: Mechanisms, Strategies, and Concepts. *Chem. Rev.* **2018**, *118*, 7409–7531.

(276) Stewart, M. P.; Sharei, A.; Ding, X.; Sahay, G.; Langer, R.; Jensen, K. F. In Vitro and Ex Vivo Strategies for Intracellular Delivery. *Nature* **2016**, *538*, 183–192.

(277) Sanchez-Lopez, C.; Labadie, N.; Lombardo, V.; Biglione, F.; Manta, B.; Jacob, R.; Gladyshev, V.; Abdelilah-Seyfried, S.; Selenko, P.; Binolfi, A. An NMR-based Biosensor to Measure Stereo-specific Methionine Sulfoxide Reductase (MSR) Activities in Vitro and in Vivo. *Chem. - Eur. J.* **2020**, chem.202002645-7.

(278) Thole, J. F.; Fadero, T. C.; Bonin, J. P.; Stadmiller, S. S.; Giudice, J. A.; Pielak, G. J. Danio Rerio Oocytes for Eukaryotic In-Cell NMR. *Biochemistry* **2021**, *60*, 451–459.

(279) Selenko, P.; Wagner, G. Looking into Live Cells with In-Cell NMR Spectroscopy. *J. Struct. Biol.* **2007**, *158*, 244–253.

(280) Sakai, T.; Tochio, H.; Tenno, T.; Ito, Y.; Kokubo, T.; Hiroaki, H.; Shirakawa, M. In-Cell NMR Spectroscopy of Proteins inside Xenopus Laevis Oocytes. *J. Biomol. NMR* **2006**, *36*, 179–188.

(281) 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.* **2006**, *103*, 11904–11909.

(282) Selenko, P.; Frueh, D. P.; Elsaesser, S. J.; Haas, W.; Gygi, S. P.; Wagner, G. In Situ Observation of Protein Phosphorylation by High-Resolution NMR Spectroscopy. *Nat. Struct. Mol. Biol.* **2008**, *15*, 321–329.

(283) Bodart, J.-F.; Wieruszeski, J.-M.; Amniai, L.; Leroy, A.; Landrieu, I.; Rousseau-Lescuyer, A.; Vilain, J.-P.; Lippens, G. NMR Observation of Tau in Xenopus Oocytes. *J. Magn. Reson.* **2008**, *192*, 252–257.

(284) Amata, I.; Maffei, M.; Igea, A.; Gay, M.; Vilaseca, M.; Nebreda, A. R.; Pons, M. Multi-Phosphorylation of the Intrinsically Disordered Unique Domain of c-Src Studied by in-Cell and Real-Time NMR Spectroscopy. *ChemBioChem* **2013**, *14*, 1820–1827.

(285) Luh, L. M.; Hänsel, R.; Löhr, F.; Kirchner, D. K.; Krauskopf, K.; Pitzius, S.; Schäfer, B.; Tufar, P.; Corbeski, I.; Güntert, P.; et al. Molecular Crowding Drives Active Pin1 into Nonspecific Complexes with Endogenous Proteins Prior to Substrate Recognition. *J. Am. Chem. Soc.* **2013**, *135*, 13796–13803.

(286) Ye, Y.; Liu, X.; Chen, Y.; Xu, G.; Wu, Q.; Zhang, Z.; Yao, C.; Liu, M.; Li, C. Labeling Strategy and Signal Broadening Mechanism of Protein NMR Spectroscopy in Xenopus Laevis Oocytes. *Chem. - Eur. J.* **2015**, *21*, 8686–8690.

(287) 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.

(288) 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.

(289) Hänsel, R.; Foldynová-Trantírková, S.; Löhr, F.; Buck, J.; Bongartz, E.; Bamberg, E.; Schwalbe, H.; Dötsch, V.; Trantírek, L. Evaluation of Parameters Critical for Observing Nucleic Acids Inside Living *Xenopus Laevis* Oocytes by In-Cell NMR Spectroscopy. *J. Am. Chem. Soc.* **2009**, *131*, 15761–15768.

(290) Hänsel, R.; Löhr, F.; Foldynová-Trantírková, S.; Bamberg, E.; Trantírek, L.; Dötsch, V. The Parallel G-Quadruplex Structure of Vertebrate Telomeric Repeat Sequences Is Not the Preferred Folding Topology under Physiological Conditions. *Nucleic Acids Res.* **2011**, *39*, 5768–5775.

(291) Salgado, G. F.; Cazenave, C.; Kerkour, A.; Mergny, J.-L. G-Quadruplex DNA and Ligand Interaction in Living Cells Using NMR Spectroscopy. *Chem. Sci.* **2015**, *6*, 3314–3320.

(292) Bao, H.-L.; Ishizuka, T.; Sakamoto, T.; Fujimoto, K.; Uechi, T.; Kenmochi, N.; Xu, Y. Characterization of Human Telomere RNA G-Quadruplex Structures in Vitro and in Living Cells Using 19F NMR Spectroscopy. *Nucleic Acids Res.* **2017**, *45*, 5501–5511.

(293) Manna, S.; Sarkar, D.; Srivatsan, S. G. A Dual-App Nucleoside Probe Provides Structural Insights into the Human Telomeric Overhang in Live Cells. *J. Am. Chem. Soc.* **2018**, *140*, 12622–12633.

(294) Broft, P.; Dzatko, S.; Krafcikova, M.; Wacker, A.; Hansel-Hertsch, R.; Dötsch, V.; Trantirek, L.; Schwalbe, H. In-Cell NMR Spectroscopy of Functional Riboswitch Aptamers in Eukaryotic Cells. *Angew. Chem.* **2020**, *15*, 679–10.

(295) Thongwichian, R.; Selenko, P. In-Cell NMR in Xenopus Laevis Oocytes. In *Intrinsically Disordered Protein Analysis: Volume 1, Methods and Experimental Tools*; Uversky, V. N., Dunker, A. K., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; pp 33–41.

(296) Azarkh, M.; Singh, V.; Okle, O.; Seemann, I. T.; Dietrich, D. R.; Hartig, J. S.; Drescher, M. Site-Directed Spin-Labeling of Nucleotides and the Use of in-Cell EPR to Determine Long-Range Distances in a Biologically Relevant Environment. *Nat. Protoc.* **2013**, *8*, 131–147.

(297) Krafcikova, M.; Hänsel-Hertsch, R.; Trantirek, L.; Foldynova-Trantirkova, S. In Cell NMR Spectroscopy: Investigation of G-Quadruplex Structures Inside Living Xenopus Laevis Oocytes. In *G-Quadruplex Nucleic Acids*; Yang, D., Lin, C., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2019; Vol. 2035, pp 397–405.

(298) Thongwichian, R.; Kosten, J.; Benary, U.; Rose, H. M.; Stuiver, M.; Theillet, F.-X.; Dose, A.; Koch, B.; Yokoyama, H.; Schwarzer, D.; et al. A Multiplexed NMR-Reporter Approach to Measure Cellular Kinase and Phosphatase Activities in Real-Time. *J. Am. Chem. Soc.* **2015**, *137*, 6468–6471.

(299) Barraud, P.; Gato, A.; Heiss, M.; Catala, M.; Kellner, S.; Tisné, C. Time-Resolved NMR Monitoring of TRNA Maturation. *Nat. Commun.* **2019**, 1–14.

(300) Kauffman, W. B.; Fuselier, T.; He, J.; Wimley, W. C. Mechanism Matters: A Taxonomy of Cell Penetrating Peptides. *Trends Biochem. Sci.* **2015**, *40*, 749–764.

(301) Guidotti, G.; Brambilla, L.; Rossi, D. Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends Pharmacol. Sci.* **2017**, *38*, 406–424.

(302) Peraro, L.; Kritzer, J. A. Emerging Methods and Design Principles for Cell-Penetrant Peptides. *Angew. Chem. Int. Ed.* **2018**, *57*, 11868–11881.

(303) Dougherty, P. G.; Sahni, A.; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **2019**, *119*, 10241–10287.

(304) Lee, H.-M.; Ren, J.; Tran, K. M.; Jeon, B.-M.; Park, W.-U.; Kim, H.; Lee, K. E.; Oh, Y.; Choi, M.; Kim, D.-S.; et al. Identification of Efficient Prokaryotic Cell-Penetrating Peptides with Applications in Bacterial Biotechnology. *Commun. Biol.* **2021**, *4*, 205.

(305) Hoyer, J.; Neundorf, I. Peptide Vectors for the Nonviral Delivery of Nucleic Acids. *Acc. Chem. Res.* **2012**, *45*, 1048–1056.

(306) Ida, H.; Takahashi, Y.; Kumatani, A.; Shiku, H.; Murayama, T.; Hirose, H.; Futaki, S.; Matsue, T. Nanoscale Visualization of Morphological Alteration of Live-Cell Membranes by the Interaction with Oligoarginine Cell-Penetrating Peptides. *Anal. Chem.* **2021**, *93*, 5383–5393.

(307) Deprey, K.; Becker, L.; Kritzer, J.; Plückthun, A. Trapped! A Critical Evaluation of Methods for Measuring Total Cellular Uptake versus Cytosolic Localization. *Bioconjug. Chem.* **2019**, *30*, 1006–1027.

(308) Qian, Z.; Martyna, A.; Hard, R. L.; Wang, J.; Appiah-Kubi, G.; Coss, C.; Phelps, M. A.; Rossman, J. S.; Pei, D. Discovery and Mechanism of Highly Efficient Cyclic Cell-Penetrating Peptides. *Biochemistry* **2016**, *55*, 2601–2612.

(309) Futaki, S. Functional Peptides That Target Biomembranes: Design and Modes of Action. *Chem. Pharm. Bull. (Tokyo)* **2021**, *69*, 601–607.

(310) Inomata, K.; Ohno, A.; Tochio, H.; Isogai, S.; Tenno, T.; Nakase, I.; Takeuchi, T.; Futaki, S.; Ito, Y.; Hiroaki, H.; et al. High-Resolution Multi-Dimensional NMR Spectroscopy of Proteins in Human Cells. *Nature* **2009**, *457*, 106–109.

(311) Danielsson, J.; Inomata, K.; Murayama, S.; Tochio, H.; Lang, L.; Shirakawa, M.; Oliveberg, M. Pruning the ALS-Associated Protein SOD1 for in-Cell NMR. *J. Am. Chem. Soc.* **2013**, *135*, 10266–10269.

(312) Hembram, D. S. S.; Haremaki, T.; Hamatsu, J.; Inoue, J.; Kamoshida, H.; Ikeya, T.; Mishima, M.; Mikawa, T.; Hayashi, N.; Shirakawa, M.; et al. An In-Cell NMR Study of Monitoring Stress-Induced Increase of Cytosolic Ca2+ Concentration in HeLa Cells. *Biochem. Biophys. Res. Commun.* **2013**, *438*, 653–659.

(313) Takeuchi, T.; Kosuge, M.; Tadokoro, A.; Sugiura, Y.; Nishi, M.; Kawata, M.; Sakai, N.; Matile, S.; Futaki, S. Direct and Rapid Cytosolic Delivery Using Cell-Penetrating Peptides Mediated by Pyrenebutyrate. *ACS Chem. Biol.* **2006**, *1*, 299–303.

(314) Inomata, K.; Kigawa, T.; Shirakawa, M. *In-Cell NMR in Eukaryotic Cells – I (Introduction from Outside of Cells)*; New Developments in NMR; 2020; Vol. 2020-January, p 44.

(315) Lv, S.; Sylvestre, M.; Prossnitz, A. N.; Yang, L. F.; Pun, S. H. Design of Polymeric Carriers for Intracellular Peptide Delivery in Oncology Applications. *Chem. Rev.* **2021**, acs.chemrev.0c00963.

(316) Yu, S.; Yang, H.; Li, T.; Pan, H.; Ren, S.; Luo, G.; Jiang, J.; Yu, L.; Chen, B.; Zhang, Y.; et al. Efficient Intracellular Delivery of Proteins by a Multifunctional Chimaeric Peptide in Vitro and in Vivo. *Nat. Commun.* **2021**, *12*, 5131.

(317) Peraro, M. D.; van der Goot, F. G. Pore-Forming Toxins: Ancient, but Never Really out of Fashion. *Nat. Rev. Microbiol.* **2016**, *14*, 77–92.

(318) van Pee, K.; Neuhaus, A.; D'Imprima, E.; Mills, D. J.; Kühlbrandt, W.; Yildiz, Ö. CryoEM Structures of Membrane Pore and Prepore Complex Reveal Cytolytic Mechanism of Pneumolysin. *eLife* **2017**, *6*, e23644.

(319) Andrews, N. W.; Almeida, P. E.; Corrotte, M. Damage Control: Cellular Mechanisms of Plasma Membrane Repair. *Trends Cell Biol.* **2014**, *24*, 734–742.

(320) Zhen, Y.; Radulovic, M.; Vietri, M.; Stenmark, H. Sealing Holes in Cellular Membranes. *EMBO J.* **2021**, *40*.

(321) Babiychuk, E. B.; Monastyrskaya, K.; Potez, S.; Draeger, A. Intracellular Ca2+ Operates a Switch between Repair and Lysis of Streptolysin O-Perforated Cells. *Cell Death Differ.* **2009**, *16*, 1126–1134.

(322) Cooper, S. T.; McNeil, P. L. Membrane Repair: Mechanisms and Pathophysiology. *Physiol. Rev.* **2015**, *95*, 1205–1240.

(323) Wolfmeier, H.; Schoenauer, R.; Atanassoff, A. P.; Neill, D. R.; Kadioglu, A.; Draeger, A.; Babiychuk, E. B. Ca2+-Dependent Repair of Pneumolysin Pores: A New Paradigm for Host Cellular Defense against Bacterial Pore-Forming Toxins. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2015**, *1853*, 2045–2054.

(324) Ogino, S.; Kubo, S.; Umemoto, 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.

(325) Yamaoki, Y.; Kiyoishi, A.; Miyake, M.; Kano, F.; Murata, M.; Nagata, T.; Katahira, M. The First Successful Observation of In-Cell NMR Signals of DNA and RNA in Living Human Cells. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2982–2985.

(326) 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.

(327) Zhao, Q.; Fujimiya, R.; Kubo, S.; Marshall, C. B.; Ikura, M.; Shimada, I.; Nishida, N. Real-Time In-Cell NMR Reveals the Intracellular Modulation of GTP-Bound Levels of RAS. *Cell Rep.* **2020**, *32*, 108074.

(328) Sakamoto, T.; Yamaoki, Y.; Nagata, T.; Katahira, M. Detection of Parallel and Antiparallel DNA Triplex Structures in Living Human Cells Using In-Cell NMR. *Chem. Commun.* **2021**, *57*, 6364–6367.

(329) Bao, H.-L.; Liu, H.; Xu, Y. Hybrid-Type and Two-Tetrad Antiparallel Telomere DNA G-Quadruplex Structures in Living Human Cells. *Nucleic Acids Res.* **2019**, *47*, 4940–4947.

(330) Chau, C.; Actis, P.; Hewitt, E. Methods for Protein Delivery into Cells: From Current Approaches to Future Perspectives. *Biochem. Soc. Trans.* **2020**, *48*, 357–365.

(331) Binolfi, A.; Limatola, A.; Verzini, S.; Kosten, J.; Theillet, F.-X.; May Rose, H.; Bekei, B.; Stuiver, 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.

(332) Rols, M. P.; Teissié, J. Electropermeabilization of Mammalian Cells. Quantitative Analysis of the Phenomenon. *Biophys. J.* **1990**, *58*, 1089–1098.

(333) Rols, M.-P.; Teissié, J. Electropermeabilization of Mammalian Cells to Macromolecules: Control by Pulse Duration. *Biophys. J.* **1998**, *75*, 1415–1423.

(334) Guo, H.; Hao, R.; Wei, Y.; Sun, D.; Sun, S.; Zhang, Z. Optimization of Electrotransfection Conditions of Mammalian Cells with Different Biological Features. *J. Membr. Biol.* **2012**, *245*, 789–795.

(335) Li, J.; Tan, W.; Yu, M.; Lin, H. The Effect of Extracellular Conductivity on Electroporation-Mediated Molecular Delivery. *Biochim. Biophys. Acta BBA - Biomembr.* **2013**, *1828*, 461–470.

(336) Kotnik, T.; Rems, L.; Tarek, M.; Miklavčič, D. Membrane Electroporation and Electropermeabilization: Mechanisms and Models. *Annu. Rev. Biophys.* **2019**, *48*, 63–91.

(337) Muralidharan, A.; Rems, L.; Kreutzer, M. T.; Boukany, P. E. Actin Networks Regulate the Cell Membrane Permeability during Electroporation. *Biochim. Biophys. Acta BBA - Biomembr.* **2021**, *1863*, 183468.

(338) Gabriel, B.; Teissié, J. Control by Electrical Parameters of Short- and Long-Term Cell Death Resulting from Electropermeabilization of Chinese Hamster Ovary Cells. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1995**, *1266*, 171–178.

(339) Shirakashi, R.; Sukhorukov, V. L.; Tanasawa, I.; Zimmermann, U. Measurement of the Permeability and Resealing Time Constant of the Electroporated Mammalian Cell Membranes. *Int. J. Heat Mass Transf.* **2004**, *47*, 4517–4524.

(340) Sengel, J. T.; Wallace, M. I. Imaging the Dynamics of Individual Electropores. *Proc. Natl. Acad. Sci.* **2016**, *113*, 5281–5286.

(341) Krassen, H.; Pliquett, U.; Neumann, E. Nonlinear Current–Voltage Relationship of the Plasma Membrane of Single CHO Cells. *Bioelectrochemistry* **2007**, *70*, 71–77.

(342) Batista Napotnik, T.; Polajžer, T.; Miklavčič, D. Cell Death Due to Electroporation – A Review. *Bioelectrochemistry* **2021**, *141*, 107871.

(343) Sachdev, S.; Feijoo Moreira, S.; Keehnen, Y.; Rems, L.; Kreutzer, M. T.; Boukany, P. E. DNA-Membrane Complex Formation during Electroporation Is DNA Size-Dependent. *Biochim. Biophys. Acta BBA - Biomembr.* **2020**, *1862*, 183089.

(344) Kandušer, M.; Šentjurc, M.; Miklavčič, D. The Temperature Effect during Pulse Application on Cell Membrane Fluidity and Permeabilization. *Bioelectrochemistry* **2008**, *74*, 52–57.

(345) Jakstys, B.; Jakutaviciute, M.; Uzdavinyte, D.; Satkauskiene, I.; Satkauskas, S. Correlation between the Loss of Intracellular Molecules and Cell Viability after Cell Electroporation. *Bioelectrochemistry* **2020**, *135*, 107550.

(346) Wang, T.-Y.; Libardo, M. D. J.; Angeles-Boza, A. M.; Pellois, J.-P. Membrane Oxidation in Cell Delivery and Cell Killing Applications. *ACS Chem. Biol.* **2017**, *12*, 1170–1182.

(347) Phez, E.; Gibot, L.; Rols, M.-P. How Transient Alterations of Organelles in Mammalian Cells Submitted to Electric Field May Explain Some Aspects of Gene Electrotransfer Process. *Bioelectrochemistry* **2016**, *112*, 166–172.

(348) van den Hoff, M. J. B.; Moorman, A. F. M.; Lamers, W. H. Electroporation in 'Intracellular' Buffer Increases Cell Survival. *Nucleic Acids Res.* **1992**, *20*, 2902–2902.

(349) Sherba, J. J.; Hogquist, S.; Lin, H.; Shan, J. W.; Shreiber, D. I.; Zahn, J. D. The Effects of Electroporation Buffer Composition on Cell Viability and Electro-Transfection Efficiency. *Sci. Rep.* **2020**, *10*, 3053.

(350) Sadik, M. M.; Yu, M.; Zheng, M.; Zahn, J. D.; Shan, J. W.; Shreiber, D. I.; Lin, H. Scaling Relationship and Optimization of Double-Pulse Electroporation. *Biophys. J.* **2014**, *106*, 801–812.

(351) Demiryurek, Y.; Nickaeen, M.; Zheng, M.; Yu, M.; Zahn, J. D.; Shreiber, D. I.; Lin, H.; Shan, J. W. Transport, Resealing, and Re-Poration Dynamics of Two-Pulse Electroporation-Mediated Molecular Delivery. *Biochim. Biophys. Acta BBA - Biomembr.* **2015**, *1848*, 1706–1714.

(352) Pliquett, U.; Gift, E. A.; Weaver, J. C. Determination of the Electric Field and Anomalous Heating Caused by Exponential Pulses with Aluminum Electrodes in Electroporation Experiments. *Bioelectrochem. Bioenerg.* **1996**, *39*, 39–53.

(353) Kurata, K.; Yoshii, T.; Uchida, S.; Fukunaga, T.; Takamatsu, H. Visualization of Electroporation-Induced Temperature Rise Using Temperature-Sensitive Ink. *Int. J. Heat Mass Transf.* **2012**, *55*, 7207–7212.

(354) Li, Y.; Wu, M.; Zhao, D.; Wei, Z.; Zhong, W.; Wang, X.; Liang, Z.; Li, Z. Electroporation on Microchips: The Harmful Effects of PH Changes and Scaling Down. *Sci. Rep.* **2016**, *5*, 17817.

(355) Kotnik, T.; Pucihar, G.; Miklavčič, D. Induced Transmembrane Voltage and Its Correlation with Electroporation-Mediated Molecular Transport. *J. Membr. Biol.* **2010**, *236*, 3–13.

(356) Yarmush, M. L.; Golberg, A.; Serša, G.; Kotnik, T.; Miklavčič, D. Electroporation-Based Technologies for Medicine: Principles, Applications, and Challenges. *Annu. Rev. Biomed. Eng.* **2014**, *16*, 295–320.

(357) Limatola, A.; Eichmann, C.; Jacob, R. S.; Ben-Nissan, G.; Sharon, M.; Binolfi, A.; Selenko, P. Time-Resolved NMR Analysis of Proteolytic α-Synuclein Processing in Vitro and in Cellulo. *PROTEOMICS* **2018**, *18*, 1800056–13.

(358) Danielsson, J.; Mu, X.; Lang, L.; Wang, H.; Binolfi, A.; Theillet, F.-X.; Bekei, B.; Logan, D. T.; Selenko, P.; Wennerstrom, H.; et al. Thermodynamics of Protein Destabilization in Live Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 12402–12407.

(359) 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.

(360) Burmann, B. M.; Gerez, J. A.; Matečko-Burmann, I.; Campioni, S.; Kumari, P.; Ghosh, D.; Mazur, A.; Aspholm, E. E.; Šulskis, D.; Wawrzyniuk, M.; et al. Regulation of α-Synuclein by Chaperones in Mammalian Cells. *Nature* **2019**, 1–25.

(361) Schlagnitweit, J.; Friebe Sandoz, S.; Jaworski, A.; Guzzetti, I.; Aussenac, F.; Carbajo, R. J.; Chiarparin, E.; Pell, A. J.; Petzold, K. Observing an Antisense Drug Complex in Intact Human Cells by In-Cell NMR Spectroscopy. *ChemBioChem* **2019**, *20*, 2474–2478.

(362) Hikone, Y.; Hirai, G.; Mishima, M.; Inomata, K.; Ikeya, T.; Arai, S.; Shirakawa, M.; Sodeoka, M.; Ito, Y. A New Carbamidemethyl-Linked Lanthanoid Chelating Tag for PCS NMR Spectroscopy of Proteins in Living HeLa Cells. *J. Biomol. NMR* **2016**, *66*, 99–110.

(363) Inomata, K.; Kamoshida, H.; Ikari, M.; Ito, Y.; Kigawa, T. Impact of Cellular Health Conditions on the Protein Folding State in Mammalian Cells. *Chem. Commun.* **2017**, *53*, 11245–11248.

(364) Viskova, P.; Krafcik, D.; Trantirek, L.; Foldynova-Trantirkova, S. In-Cell NMR Spectroscopy of Nucleic Acids in Human Cells. *Curr. Protoc. Nucleic Acid Chem.* **2019**, *76*, e71.

(365) Krafcikova, M.; Dzatko, S.; Caron, C.; Granzhan, A.; Fiala, R.; Loja, T.; Teulade-Fichou, M.-P.; Fessl, T.; Hänsel-Hertsch, R.; Mergny, J.-L.; et al. Monitoring DNA–Ligand Interactions in Living Human Cells Using NMR Spectroscopy. *J. Am. Chem. Soc.* **2019**, *141*, 13281–13285.

(366) Leeb, S.; Yang, F.; Oliveberg, M.; Danielsson, J. Connecting Longitudinal and Transverse Relaxation Rates in Live-Cell NMR. *J. Phys. Chem. B* **2020**, *124*, 10698–10707.

(367) Zhang, S.; Wang, C.; Lu, J.; Ma, X.; Liu, Z.; Li, D.; Liu, Z.; Liu, C. In-Cell NMR Study of Tau and MARK2 Phosphorylated Tau. *Int. J. Mol. Sci.* **2019**, *20*, 90–14.

(368) De Biase, P. M.; Paggi, D. A.; Doctorovich, F.; Hildebrandt, P.; Estrin, D. A.; Murgida, D. H.; Marti, M. A. Molecular Basis for the Electric Field Modulation of Cytochrome *c* Structure and Function. *J. Am. Chem. Soc.* **2009**, *131*, 16248–16256.

(369) Marracino, P.; Apollonio, F.; Liberti, M.; d'Inzeo, G.; Amadei, A. Effect of High Exogenous Electric Pulses on Protein Conformation: Myoglobin as a Case Study. *J. Phys. Chem. B* **2013**, *117*, 2273–2279.

(370) Wang, X.; Li, Y.; He, X.; Chen, S.; Zhang, J. Z. H. Effect of Strong Electric Field on the Conformational Integrity of Insulin. *J. Phys. Chem. A* **2014**, *118*, 8942–8952.

(371) della Valle, E.; Marracino, P.; Pakhomova, O.; Liberti, M.; Apollonio, F. Nanosecond Pulsed Electric Signals Can Affect Electrostatic Environment of Proteins below the Threshold of Conformational Effects: The Case Study of SOD1 with a Molecular Simulation Study. *PLOS ONE* **2019**, *14*, e0221685.

(372) Jiang, Z.; You, L.; Dou, W.; Sun, T.; Xu, P. Effects of an Electric Field on the Conformational Transition of the Protein: A Molecular Dynamics Simulation Study. *Polymers* **2019**, *11*, 282.

(373) Arbeitman, C. R.; Rojas, P.; Ojeda-May, P.; Garcia, M. E. The SARS-CoV-2 Spike Protein Is Vulnerable to Moderate Electric Fields. *Nat. Commun.* **2021**, *12*, 5407.

(374) Urabe, G.; Sato, T.; Nakamura, G.; Kobashigawa, Y.; Morioka, H.; Katsuki, S. 1.2 MV/Cm Pulsed Electric Fields Promote Transthyretin Aggregate Degradation. *Sci. Rep.* **2020**, *10*, 12003.

(375) Marracino, P.; Havelka, D.; Průša, J.; Liberti, M.; Tuszynski, J.; Ayoub, A. T.; Apollonio, F.; Cifra, M. Tubulin Response to Intense Nanosecond-Scale Electric Field in Molecular Dynamics Simulation. *Sci. Rep.* **2019**, *9*, 10477.

(376) Hekstra, D. R.; White, K. I.; Socolich, M. A.; Henning, R. W.; Šrajer, V.; Ranganathan, R. Electric-Field-Stimulated Protein Mechanics. *Nature* **2016**, *540*, 400–405.

(377) Zhao, W.; Yang, R. The Effect of Pulsed Electric Fields on the Inactivation and Structure of Lysozyme. *Food Chem.* **2008**, *110*, 334–343.

(378) Zhao, W.; Yang, R. Experimental Study on Conformational Changes of Lysozyme in Solution Induced by Pulsed Electric Field and Thermal Stresses. *J. Phys. Chem. B* **2010**, *114*, 503–510.

(379) Beebe, S. J. Considering Effects of Nanosecond Pulsed Electric Fields on Proteins. *Bioelectrochemistry* **2015**, *103*, 52–59.

(380) Wu, L.; Zhao, W.; Yang, R.; Chen, X. Effects of Pulsed Electric Fields Processing on Stability of Egg White Proteins. *J. Food Eng.* **2014**, *139*, 13–18.

(381) Urabe, G.; Katagiri, T.; Katsuki, S. Intense Pulsed Electric Fields Denature Urease Protein. *Bioelectricity* **2020**, *2*, 33–39.

(382) Bekard, I.; Dunstan, D. E. Electric Field Induced Changes in Protein Conformation. *Soft Matter* **2014**, *10*, 431–437.

(383) Shi, Y.; Acerson, M. J.; Shuford, K. L.; Shaw, B. F. Voltage-Induced Misfolding of Zinc-Replete ALS Mutant Superoxide Dismutase-1. *ACS Chem. Neurosci.* **2015**, *6*, 1696–1707.

(384) Chakrabarti, R.; Wylie, D. E.; Schuster, S. M. Transfer of Monoclonal Antibodies into Mammalian Cells by Electroporation. *J. Biol. Chem.* **1989**, *264*, 15494–15500.

(385) Lukas, J.; Bartek, J.; Strauss, M. Efficient Transfer of Antibodies into Mammalian Cells by Electroporation. *J. Immunol. Methods* **1994**, *170*, 255–259.

(386) Baron, S.; Poast, J.; Rizzo, D.; McFarland, E.; Kieff, E. Electroporation of Antibodies, DNA, and Other Macromolecules into Cells: A Highly Efficient Method. *J. Immunol. Methods* **2000**, *242*, 115–126.

(387) Freund, G.; Sibler, A.-P.; Desplancq, D.; Oulad-Abdelghani, M.; Vigneron, M.; Gannon, J.; Van Regenmortel, M. H.; Weiss, E. Targeting Endogenous Nuclear Antigens by Electrotransfer of Monoclonal Antibodies in Living Cells. *mAbs* **2013**, *5*, 518–522.

(388) Conic, S.; Desplancq, D.; Ferrand, A.; Fischer, V.; Heyer, V.; Reina San Martin, B.; Pontabry, J.; Oulad-Abdelghani, M.; Babu N., K.; Wright, G. D.; et al. Imaging of Native Transcription Factors and Histone Phosphorylation at High Resolution in Live Cells. *J. Cell Biol.* **2018**, *217*, 1537–1552.

(389) Alex, A.; Piano, V.; Polley, S.; Stuiver, M.; Voss, S.; Ciossani, G.; Overlack, K.; Voss, B.; Wohlgemuth, S.; Petrovic, A.; et al. Electroporated Recombinant Proteins as Tools for in Vivo Functional Complementation, Imaging and Chemical Biology. *eLife* **2019**, *8*, 2026.

(390) Yang, Y.; Yang, F.; Gong, Y.-J.; Chen, J.-L.; Goldfarb, D.; Su, X.-C. A Reactive, Rigid Gd ^{III} Labeling Tag for In-Cell EPR Distance Measurements in Proteins. *Angew. Chem. Int. Ed.* **2017**, *56*, 2914–2918.

(391) Yang, Y.; Yang, F.; Li, X.-Y.; Su, X.-C.; Goldfarb, D. In-Cell EPR Distance Measurements on Ubiquitin Labeled with a Rigid PyMTA-Gd(III) Tag. *J. Phys. Chem. B* **2019**, *123*, 1050–1059.

(392) Yang, Y.; Pan, B.-B.; Tan, X.; Yang, F.; Liu, Y.; Su, X.-C.; Goldfarb, D. In-Cell Trityl-Trityl Distance Measurements on Proteins. *J. Phys. Chem. Lett.* **2020**, *11*, 1141–1147.

(393) Yang, Y.; Chen, S.-N.; Yang, F.; Li, X.-Y.; Feintuch, A.; Su, X.-C.; Goldfarb, D. In-Cell Destabilization of a Homodimeric Protein Complex Detected by DEER Spectroscopy. *Proc. Natl. Acad. Sci.* **2020**, *117*, 20566–20575.

(394) Kucher, S.; Elsner, C.; Safonova, M.; Maffini, S.; Bordignon, E. In-Cell Double Electron–Electron Resonance at Nanomolar Protein Concentrations. *J. Phys. Chem. Lett.* **2021**, *12*, 3679–3684.

(395) 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.

(396) Breindel, L.; Burz, D. S.; Shekhtman, A. Active Metabolism Unmasks Functional Protein–Protein Interactions in Real Time in-Cell NMR. *Commun. Biol.* **2020**, *3*, 1–9.

(397) Luchinat, E.; Barbieri, L.; Cremonini, M.; Pennestri, M.; Nocentini, A.; Supuran, C. T.; Banci, L. Determination of Intracellular Protein–Ligand Binding Affinity by Competition Binding in-Cell NMR. *Acta Crystallogr. Sect. Struct. Biol.* **2021**, *77*, 1270–1281.

(398) Mateos, B.; Sealey Cardona, M.; Balazs, K.; Konrat, J.; Staffler, G.; Konrat, R. NMR Characterization of Surface Receptor Protein Interactions in Live Cells Using Methylcellulose Hydrogels. *Angew. Chem.* **2020**, *59*, 3886–3890.

(399) Alshamleh, I.; Krause, N.; Richter, C.; Kurrle, N.; Serve, H.; Günther, U. L.; Schwalbe, H. Real-Time NMR Spectroscopy for Studying Metabolism. *Angew. Chem. Int. Ed Engl.* **2020**, *59*, 2304–2308.

(400) Theillet, F.-X.; Smet-Nocca, C.; Liokatis, S.; Thongwichian, R.; Kosten, J.; Yoon, M.-K.; Kriwacki, R. W.; Landrieu, I.; Lippens, G.; Selenko, P. Cell Signaling, Post-Translational Protein Modifications and NMR Spectroscopy. *J. Biomol. NMR* **2012**, *54*, 217–236.

(401) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(402) Kitevski-LeBlanc, J. L.; Prosser, R. S. Current Applications of 19F NMR to Studies of Protein Structure and Dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1–33.

(403) Tamiola, K.; Mulder, F. A. A. Using NMR Chemical Shifts to Calculate the Propensity for Structural Order and Disorder in Proteins. *Biochem. Soc. Trans.* **2012**, *40*, 1014–1020.

(404) Nielsen, J. T.; Mulder, F. A. A. POTENCI: Prediction of Temperature, Neighbor and PH-Corrected Chemical Shifts for Intrinsically Disordered Proteins. *J. Biomol. NMR* **2018**, *70*, 141–165.

(405) Conibear, A. C.; Rosengren, K. J.; Becker, C. F. W.; Kaehlig, H. Random Coil Shifts of Posttranslationally Modified Amino Acids. *J. Biomol. NMR* **2019**, *73*, 587–599.

(406) Hendus-Altenburger, R.; Fernandes, C. B.; Bugge, K.; Kunze, M. B. A.; Boomsma, W.; Kragelund, B. B. Random Coil Chemical Shifts for Serine, Threonine and Tyrosine Phosphorylation over a Broad PH Range. *J. Biomol. NMR* **2019**, *73*, 713–725.

(407) Kumar, A.; Narayanan, V.; Sekhar, A. Characterizing Post-Translational Modifications and Their Effects on Protein Conformation Using NMR Spectroscopy. *Biochemistry* **2019**, *59*, 57–73.

(408) Hafsa, N. E.; Berjanskii, M. V.; Arndt, D.; Wishart, D. S. Rapid and Reliable Protein Structure Determination via Chemical Shift Threading. *J. Biomol. NMR* **2018**, *70*, 33–51.

(409) Nerli, S.; McShan, A. C.; Sgourakis, N. G. Chemical Shift-Based Methods in NMR Structure Determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **2018**, *106–107*, 1–25.

(410) Boeszoermenyi, A.; Ogórek, B.; Jain, A.; Arthanari, H.; Wagner, G. The Precious Fluorine on the Ring: Fluorine NMR for Biological Systems. *J. Biomol. NMR* **2020**, *74*, 365–379.

(411) Schanda, P.; Kupče, Ē.; Brutscher, B. SOFAST-HMQC Experiments for Recording Two-Dimensional Heteronuclear Correlation Spectra of Proteins within a Few Seconds. *J. Biomol. NMR* **2005**, *33*, 199–211.

(412) Paul Schanda. Fast-Pulsing Longitudinal Relaxation Optimized Techniques: Enriching the Toolbox of Fast Biomolecular NMR Spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *55*, 238–265.

(413) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. Attenuated T2 Relaxation by Mutual Cancellation of Dipole-Dipole Coupling and Chemical Shift Anisotropy Indicates an Avenue to NMR Structures of Very Large Biological Macromolecules in Solution. *Proc. Natl. Acad. Sci.* **1997**, *94*, 12366–12371.

(414) Pervushin, K.; Vögeli, B.; Eletsky, A. Longitudinal 1H Relaxation Optimization in TROSY NMR Spectroscopy. *J. Am. Chem. Soc.* **2002**, *124*, 12898–12902.

(415) Mandal, P. K.; Majumdar, A. A Comprehensive Discussion of HSQC and HMQC Pulse Sequences. *Concepts Magn. Reson.* **2004**, *20A*, 1–23.

(416) Lescop, E.; Kern, T.; Brutscher, B. Guidelines for the Use of Band-Selective Radiofrequency Pulses in Hetero-Nuclear NMR: Example of Longitudinal-Relaxation-Enhanced BEST-Type 1H–15N Correlation Experiments. *J. Magn. Reson.* **2010**, *203*, 190–198.

(417) Tzakos, A. G.; Grace, C. R. R.; Lukavsky, P. J.; Riek, R. NMR Techniques for Very Large Proteins and Rnas in Solution. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 319–342.

(418) Farjon, J.; Boisbouvier, J.; Schanda, P.; Pardi, A.; Simorre, J.-P.; Brutscher, B. Longitudinal-Relaxation-Enhanced NMR Experiments for the Study of Nucleic Acids in Solution. *J. Am. Chem. Soc.* **2009**, *131*, 8571–8577.

(419) 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.* **2006**, *103*, 11904–11909.

(420) Sakai, T.; Tochio, H.; Inomata, K.; Sasaki, Y.; Tenno, T.; Tanaka, T.; Kokubo, T.; Hiroaki, H.; Shirakawa, M. Fluoroscopic Assessment of Protein Leakage during Xenopus Oocytes In-Cell NMR Experiment by Co-Injected EGFP. *Anal. Biochem.* **2007**, *371*, 247–249.

(421) Augustus, A. M.; Reardon, P. N.; Spicer, L. D. MetJ Repressor Interactions with DNA Probed by In-Cell NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5065–5069.

(422) DeMott, C. M.; Majumder, S.; Burz, D. S.; Reverdatto, S.; Shekhtman, A. Ribosome Mediated Quinary Interactions Modulate In-Cell Protein Activities. *Biochemistry* **2017**, *56*, 4117–4126.

(423) Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C.; Losacco, M.; Natile, G. Probing the Interaction of Cisplatin with the Human Copper Chaperone Atox1 by Solution and In-Cell NMR Spectroscopy. *J. Am. Chem. Soc.* **2011**, *133*, 18361–18369.

(424) Mu, X.; Choi, S.; Lang, L.; Mowray, D.; Dokholyan, N. V.; Danielsson, J.; Oliveberg, M. Physicochemical Code for Quinary Protein Interactions in Escherichia Coli. *Proc. Natl. Acad. Sci.* **2017**, *114*, E4556–E4563.

(425) Leeb, S.; Sörensen, T.; Yang, F.; Mu, X.; Oliveberg, M.; Danielsson, J. Diffusive Protein Interactions in Human versus Bacterial Cells. *Curr. Res. Struct. Biol.* **2020**, *2*, 68–78.

(426) Wall, K. P.; Hough, L. E. In-Cell NMR within Budding Yeast Reveals Cytoplasmic Masking of Hydrophobic Residues of FG Repeats. *Biophys. J.* **2018**, *115*, 1690–1695.

(427) Sörensen, T.; Leeb, S.; Danielsson, J.; Oliveberg, M. Polyanions Cause Protein Destabilization Similar to That in Live Cells. *Biochemistry* **2021**, *60*, 735–746.

(428) Theillet, F.-X.; Binolfi, A.; Liokatis, S.; Verzini, S.; Selenko, P. Paramagnetic Relaxation Enhancement to Improve Sensitivity of Fast NMR Methods: Application to Intrinsically Disordered Proteins. *J. Biomol. NMR* **2011**, *51*, 487–495.

(429) Takeuchi, K.; Arthanari, H.; Wagner, G. Perspective: Revisiting the Field Dependence of TROSY Sensitivity. *J. Biomol. NMR* **2016**, *66*, 221–225.

(430) Amero, C.; Schanda, P.; Durá, M. A.; Ayala, I.; Marion, D.; Franzetti, B.; Brutscher, B.; Boisbouvier, J. Fast Two-Dimensional NMR Spectroscopy of High Molecular Weight Protein Assemblies. *J. Am. Chem. Soc.* **2009**, *131*, 3448–3449.

(431) Sattler, M.; Schleucher, J.; Griesinger, C. Heteronuclear Multidimensional NMR Experiments for the Structure Determination of Proteins in Solution Employing Pulsed Field Gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93–158.

(432) Frueh, D. P. Practical Aspects of NMR Signal Assignment in Larger and Challenging Proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 47–75.

(433) Nowakowski, M.; Saxena, S.; Stanek, J.; Żerko, S.; Koźmiński, W. Applications of High Dimensionality Experiments to Biomolecular NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2015**, *90–91*, 49–73.

(434) Kazimierczuk, K.; Orekhov, V. Non-Uniform Sampling: Post-Fourier Era of NMR Data Collection and Processing. *Magn. Reson. Chem.* **2015**, *53*, 921–926.

(435) Shigemitsu, Y.; Ikeya, T.; Yamamoto, A.; Tsuchie, Y.; Mishima, M.; Smith, B. O.; Güntert, P.; Ito, Y. Evaluation of the Reliability of the Maximum Entropy Method for Reconstructing 3D and 4D NOESY-Type NMR Spectra of Proteins. *Biochem. Biophys. Res. Commun.* **2015**, *457*, 200–205.

(436) Alik, A.; Bouguechtouli, C.; Julien, M.; Bermel, W.; Ghouil, R.; Zinn-Justin, S.; Theillet, F.-X. Sensitivity-Enhanced 13

C-NMR Spectroscopy for Monitoring Multisite Phosphorylation at Physiological Temperature and PH. *Angew. Chem. Int. Ed Engl.* **2020**.

(437) Furuita, K.; Sugiki, T.; Takamuku, M.; Hattori, Y.; So, M.; Kawata, Y.; Ikegami, T.; Fujiwara, T.; Kojima, C. Sensitivity Enhancement by Sequential Data Acquisition for 13C-Direct Detection NMR. *J. Magn. Reson.* **2021**, *322*, 106878.

(438) Dubey, A.; Stoyanov, N.; Viennet, T.; Chhabra, S.; Elter, S.; Borggräfe, J.; Viegas, A.; Nowak, R. P.; Burdziev, N.; Petrov, O.; et al. Local Deuteration Enables NMR Observation of Methyl Groups in Proteins from Eukaryotic and Cell-free Expression Systems. *Angew. Chem. Int. Ed.* **2021**, anie.202016070.

(439) Rößler, P.; Mathieu, D.; Gossert, A. D. Enabling NMR Studies of High Molecular Weight Systems Without the Need for Deuteration: The XL-ALSOFAST Experiment with Delayed Decoupling. *Angew. Chem. Int. Ed.* **2020**, anie.202007715-10.

(440) Han, J.; Kiss, L.; Mei, H.; Remete, A. M.; Ponikvar-Svet, M.; Sedgwick, D. M.; Roman, R.; Fustero, S.; Moriwaki, H.; Soloshonok, V. A. Chemical Aspects of Human and Environmental Overload with Fluorine. *Chem. Rev.* **2021**, *121*, 4678–4742.

(441) Di Pietrantonio, C.; Pandey, A.; Gould, J.; Hasabnis, A.; Prosser, R. S. Understanding Protein Function Through an Ensemble Description: Characterization of Functional States by 19F NMR. In *Methods in Enzymology*; Elsevier, 2019; Vol. 615, pp 103–130.

(442) Overbeck, J. H.; Kremer, W.; Sprangers, R. A Suite of 19F Based Relaxation Dispersion Experiments to Assess Biomolecular Motions. *J. Biomol. NMR* **2020**, 1–14.

(443) Manglik, A.; Kim, T. H.; Masureel, M.; Altenbach, C.; Yang, Z.; Hilger, D.; Lerch, M. T.; Kobilka, T. S.; Thian, F. S.; Hubbell, W. L.; et al. Structural Insights into the Dynamic Process of B2-Adrenergic Receptor Signaling. *Cell* **2015**, *161*, 1101–1111.

(444) Ye, L.; Van Eps, N.; Zimmer, M.; Ernst, O. P.; Prosser, R. S. Activation of the A2A Adenosine G-Protein-Coupled Receptor by Conformational Selection. *Nature* **2016**, *533*, 265–268.

(445) Frei, J. N.; Broadhurst, R. W.; Bostock, M. J.; Solt, A.; Jones, A. J. Y.; Gabriel, F.; Tandale, A.; Shrestha, B.; Nietlispach, D. Conformational Plasticity of Ligand-Bound and Ternary GPCR Complexes Studied by 19F NMR of the Î²1-Adrenergic Receptor. *Nat. Commun.* **2020**, 1–14.

(446) Wang, X.; McFarland, A.; Madsen, J. J.; Aalo, E.; Ye, L. The Potential of 19F NMR Application in GPCR Biased Drug Discovery. *Trends Pharmacol. Sci.* **2021**, *42*, 19–30.

(447) Huang, S. K.; Pandey, A.; Tran, D. P.; Villanueva, N. L.; Kitao, A.; Sunahara, R. K.; Sljoka, A.; Prosser, R. S. Delineating the Conformational Landscape of the Adenosine A2A Receptor during G Protein Coupling. *Cell* **2021**, S0092867421002300.

(448) Li, C.; Lutz, E. A.; Slade, K. M.; Ruf, R. A. S.; Wang, G.-F.; Pielak, G. J. ^{19} F NMR Studies of α-Synuclein Conformation and Fibrillation. *Biochemistry* **2009**, *48*, 8578–8584.

(449) Ye, L.; Larda, S. T.; Frank Li, Y. F.; Manglik, A.; Prosser, R. S. A Comparison of Chemical Shift Sensitivity of Trifluoromethyl Tags: Optimizing Resolution in 19F NMR Studies of Proteins. *J. Biomol. NMR* **2015**, *62*, 97–103.

(450) Bao, H.-L.; Xu, Y. Telomeric DNA–RNA-Hybrid G-Quadruplex Exists in Environmental Conditions of HeLa Cells. *Chem. Commun.* **2020**, *56*, 6547–6550.

(451) Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I. M.; Ye, S.; Koksch, B. Fluorinated Amino Acids: Compatibility with Native Protein Structures and Effects on Protein–Protein Interactions. *Chem Soc Rev* **2012**, *41*, 2135–2171.

(452) Xiao, G.; Parsons, J. F.; Tesh, K.; Armstrong, R. N.; Gilliland, G. L. Conformational Changes in the Crystal Structure of Rat Glutathione Transferase M1-1 with Global Substitution of 3-Fluorotyrosine for Tyrosine. *J. Mol. Biol.* **1998**, *281*, 323–339.

(453) Kitevski-LeBlanc, J. L.; Evanics, F.; Scott Prosser, R. Optimizing 19F NMR Protein Spectroscopy by Fractional Biosynthetic Labeling. *J. Biomol. NMR* **2010**, *48*, 113–121.

(454) Campos-Olivas, R.; Aziz, R.; Helms, G. L.; Evans, J. N. S.; Gronenborn, A. M. Placement of ^{19} F into the Center of GB1: Effects on Structure and Stability. *FEBS Lett.* **2002**, *517*, 55–60.

(455) Wang, X.; Mercier, P.; Letourneau, P.-J.; Sykes, B. D. Effects of Phe-to-Trp Mutation and Fluorotryptophan Incorporation on the Solution Structure of Cardiac Troponin C, and Analysis of Its Suitability as a Potential Probe for in Situ NMR Studies. *Protein Sci.* **2005**, *14*, 2447–2460.

(456) Welte, H.; Zhou, T.; Mihajlenko, X.; Mayans, O.; Kovermann, M. What Does Fluorine Do to a Protein? Thermodynamic, and Highly-Resolved Structural Insights into Fluorine-Labelled Variants of the Cold Shock Protein. *Sci. Rep.* **2020**, *10*, 2640.

(457) Puffer, B.; Kreutz, C.; Rieder, U.; Ebert, M.-O.; Konrat, R.; Micura, R. 5-Fluoro Pyrimidines: Labels to Probe DNA and RNA Secondary Structures by 1D 19 F NMR Spectroscopy. *Nucleic Acids Res.* **2009**, *37*, 7728–7740.

(458) Ikeya, T.; Güntert, P.; Ito, Y. Protein Structure Determination in Living Cells. *Int. J. Mol. Sci.* **2019**, *20*, 2442.

(459) Shen, Y.; Bax, A. Protein Backbone and Sidechain Torsion Angles Predicted from NMR Chemical Shifts Using Artificial Neural Networks. *J. Biomol. NMR* **2013**, *56*, 227–241.

(460) Güntert, P.; Buchner, L. Combined Automated NOE Assignment and Structure Calculation with CYANA. *J. Biomol. NMR* **2015**, *62*, 453–471.

(461) Ikeya, T.; Ikeda, S.; Kigawa, T.; Ito, Y.; Güntert, P. Protein NMR Structure Refinement Based on Bayesian Inference. *J. Phys. Conf. Ser.* **2016**, *699*, 012005.

(462) Etezady-Esfarjani, T.; Herrmann, T.; Horst, R.; Wüthrich, K. Automated Protein NMR Structure Determination in Crude Cell-Extract. *J. Biomol. NMR* **2006**, *34*, 3–11.

(463) Tikole, S.; Jaravine, V.; Orekhov, V. Yu.; Güntert, P. Effects of NMR Spectral Resolution on Protein Structure Calculation. *PLoS ONE* **2013**, *8*, e68567.

(464) Buchner, L.; Güntert, P. Systematic Evaluation of Combined Automated NOE Assignment and Structure Calculation with CYANA. *J. Biomol. NMR* **2015**, *62*, 81–95.

(465) Otting, G. Protein NMR Using Paramagnetic Ions. *Annu. Rev. Biophys.* **2010**, *39*, 387–405.

(466) Koehler, J.; Meiler, J. Expanding the Utility of NMR Restraints with Paramagnetic Compounds: Background and Practical Aspects. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 360–389.

(467) Yagi, H.; Pilla, K. B.; Maleckis, A.; Graham, B.; Huber, T.; Otting, G. Three-Dimensional Protein Fold Determination from Backbone Amide Pseudocontact Shifts Generated by Lanthanide Tags at Multiple Sites. *Structure* **2013**, *21*, 883–890.

(468) Nitsche, C.; Otting, G. Pseudocontact Shifts in Biomolecular NMR Using Paramagnetic Metal Tags. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *98–99*, 20–49.

(469) Parigi, G.; Ravera, E.; Luchinat, C. Magnetic Susceptibility and Paramagnetism-Based NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *114–115*, 211–236.

(470) Ravera, E.; Cerofolini, L.; Fragai, M.; Parigi, G.; Luchinat, C. Characterization of Lanthanoid-Binding Proteins Using NMR Spectroscopy. In *Methods in Enzymology*; Elsevier, 2021; Vol. 651, pp 103–137.

(471) Müntener, T.; Joss, D.; Häussinger, D.; Hiller, S. Pseudocontact Shifts in Biomolecular NMR Spectroscopy. *Chem. Rev.* **2022**, acs.chemrev.1c00796.

(472) Shishmarev, D.; Otting, G. How Reliable Are Pseudocontact Shifts Induced in Proteins and Ligands by Mobile Paramagnetic Metal Tags? A Modelling Study. *J. Biomol. NMR* **2013**, *56*, 203–216.

(473) Ravera, E.; Parigi, G.; Luchinat, C. What Are the Methodological and Theoretical Prospects for Paramagnetic NMR in Structural Biology? A Glimpse into the Crystal Ball. *J. Magn. Reson.* **2019**, *306*, 173–179.

(474) Joss, D.; Häussinger, D. Design and Applications of Lanthanide Chelating Tags for Pseudocontact Shift NMR Spectroscopy with Biomacromolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, 1–29.

(475) Joss, D.; Winter, F.; Häussinger, D. A Novel, Rationally Designed Lanthanoid Chelating Tag Delivers Large Paramagnetic

Structural Restraints for Biomolecular NMR. *Chem. Commun.* **2020**, *56*, 12861–12864.

(476) Su, X.-C.; Chen, J.-L. Site-Specific Tagging of Proteins with Paramagnetic Ions for Determination of Protein Structures in Solution and in Cells. *Acc. Chem. Res.* **2019**, *52*, 1675–1686.

(477) Pilla, K. B.; Leman, J. K.; Otting, G.; Huber, T. Capturing Conformational States in Proteins Using Sparse Paramagnetic NMR Data. *PLoS ONE* **2015**, *10*, e0127053.

(478) Joss, D.; Häussinger. P4T-DOTA – a Lanthanide Chelating Tag Combining a Sterically Highly Overcrowded Backbone with a Reductively Stable Linker. *Chem. Commun.* **2019**, *55*, 10543–10546.

(479) Lescanne, M.; Skinner, S. P.; Blok, A.; Timmer, M.; Cerofolini, L.; Fragai, M.; Luchinat, C.; Ubbink, M. Methyl Group Assignment Using Pseudocontact Shifts with PARAssign. *J. Biomol. NMR* **2017**, *69*, 183–195.

(480) Flügge, F.; Peters, T. Complete Assignment of Ala, Ile, Leu, Met and Val Methyl Groups of Human Blood Group A and B Glycosyltransferases Using Lanthanide-Induced Pseudocontact Shifts and Methyl-Methyl NOESY. *J. Biomol. NMR* **2018**, *70*, 245–259.

(481) Pearce, B. J. G.; Jabar, S.; Loh, C.-T.; Szabo, M.; Graham, B.; Otting, G. Structure Restraints from Heteronuclear Pseudocontact Shifts Generated by Lanthanide Tags at Two Different Sites. *J. Biomol. NMR* **2017**, *68*, 19–32.

(482) Orton, H. W.; Huber, T.; Otting, G. Paramagpy: Software for Fitting Magnetic Susceptibility Tensors Using Paramagnetic Effects Measured in NMR Spectra. *Magn. Reson.* **2020**, *1*, 1–12.

(483) Cucuzza, S.; Güntert, P.; Plückthun, A.; Zerbe, O. An Automated Iterative Approach for Protein Structure Refinement Using Pseudocontact Shifts. *J. Biomol. NMR* **2021**, *75*, 319–334.

(484) Häussinger, D.; Huang, J.; Grzesiek, S. DOTA-M8: An Extremely Rigid, High-Affinity Lanthanide Chelating Tag for PCS NMR Spectroscopy. *J. Am. Chem. Soc.* **2009**, *131*, 14761–14767.

(485) Salmon, L.; Blackledge, M. Investigating Protein Conformational Energy Landscapes and Atomic Resolution Dynamics from NMR Dipolar Couplings: A Review. *Rep. Prog. Phys. Phys. Soc. G. B.* **2015**, *78*, 126601.

(486) Yao, L.; Ying, J.; Bax, A. Improved Accuracy of 15N–1H Scalar and Residual Dipolar Couplings from Gradient-Enhanced IPAP-HSQC Experiments on Protonated Proteins. *J. Biomol. NMR* **2009**, *43*, 161–170.

(487) Clore, G. M.; Iwahara, J. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. *Chem. Rev.* **2009**, *109*, 4108–4139.

(488) Trindade, I. B.; Invernici, M.; Cantini, F.; Louro, R. O.; Piccioli, M. PRE-driven Protein NMR Structures: An Alternative Approach in Highly Paramagnetic Systems. *FEBS J.* **2021**, *288*, 3010–3023.

(489) Gaalswyk, K.; Liu, Z.; Vogel, H. J.; MacCallum, J. L. An Integrative Approach to Determine 3D Protein Structures Using Sparse Paramagnetic NMR Data and Physical Modeling. *Front. Mol. Biosci.* **2021**, *8*, 676268.

(490) Karthikeyan, G.; Bonucci, A.; Casano, G.; Gerbaud, G.; Abel, S.; Thomé, V.; Kodjabachian, L.; Magalon, A.; Guigliarelli, B.; Belle, V.; et al. A Bioresistant Nitroxide Spin Label for In-Cell EPR Spectroscopy: In Vitro and In Oocytes Protein Structural Dynamics Studies. *Angew. Chem. Int. Ed.* **2018**, *57*, 1366–1370.

(491) Juliusson, H. Y.; Sigurdsson, S. Th. Reduction Resistant and Rigid Nitroxide Spin-Labels for DNA and RNA. *J. Org. Chem.* **2020**, *85*, 4036–4046.

(492) Battiste, J. L.; Wagner, G. Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data. *Biochemistry* **2000**, *39*, 5355–5365.

(493) Schwalbe, M.; Ozenne, V.; Bibow, S.; Jaremko, M.; Jaremko, L.; Gajda, M.; Jensen, M. R.; Biernat, J.; Becker, S.; Mandelkow, E.; et al. Predictive Atomic Resolution Descriptions of Intrinsically Disordered HTau40 and α-Synuclein in Solution from NMR and Small Angle Scattering. *Structure* **2014**, *22*, 238–249.

(494) Miao, Q.; Nitsche, C.; Orton, H.; Overhand, M.; Otting, G.; Ubbink, M. Paramagnetic Chemical Probes for Studying Biological Macromolecules. *Chem. Rev.* **2022**, acs.chemrev.1c00708.

(495) Orton, H. W.; Otting, G. Accurate Electron–Nucleus Distances from Paramagnetic Relaxation Enhancements. *J. Am. Chem. Soc.* **2018**, *140*, 7688–7697.

(496) Schnorr, K. A. Impact of Spin Label Rigidity on Extent and Accuracy of Distance Information from PRE Data. *J. Biomol. NMR* **2017**, *68*, 53–63.

(497) Mackereth, C. D.; Madl, T.; Bonnal, S.; Simon, B.; Zanier, K.; Gasch, A.; Rybin, V.; Valcárcel, J.; Sattler, M. Multi-Domain Conformational Selection Underlies Pre-MRNA Splicing Regulation by U2AF. *Nature* **2011**, *475*, 408–411.

(498) Hennig, J.; Warner, L. R.; Simon, B.; Geerlof, A.; Mackereth, C. D.; Sattler, M. Structural Analysis of Protein–RNA Complexes in Solution Using NMR Paramagnetic Relaxation Enhancements. In *Methods in Enzymology*; Elsevier, 2015; Vol. 558, pp 333–362.

(499) Delhommel, F.; Gabel, F.; Sattler, M. Current Approaches for Integrating Solution NMR Spectroscopy and Small-Angle Scattering to Study the Structure and Dynamics of Biomolecular Complexes. *J. Mol. Biol.* **2020**, *432*, 2890–2912.

(500) Waldeck, A. R.; Kuchel, P. W.; Lennon, A. J.; Chapman, B. E. NMR Diffusion Measurements to Characterise Membrane Transport and Solute Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *30*, 39–68.

(501) Stallmach, F.; Galvosas, P. Spin Echo NMR Diffusion Studies. In *Annual Reports on NMR Spectroscopy*; Elsevier, 2007; Vol. 61, pp 51–131.

(502) Pagès, G.; Gilard, V.; Martino, R.; Malet-Martino, M. Pulsed-Field Gradient Nuclear Magnetic Resonance Measurements (PFG NMR) for Diffusion Ordered Spectroscopy (DOSY) Mapping. *Analyst* **2017**, *142*, 3771–3796.

(503) Rajagopalan, S.; Chow, C.; Raghunathan, V.; Fry, C. G.; Cavagnero, S. NMR Spectroscopic Filtration of Polypeptides and Proteins in Complex Mixtures. *J. Biomol. NMR* **2004**, *29*, 505–516.

(504) Augustyniak, R.; Ferrage, F.; Paquin, R.; Lequin, O.; Bodenhausen, G. Methods to Determine Slow Diffusion Coefficients of Biomolecules. Applications to Engrailed 2, a Partially Disordered Protein. *J. Biomol. NMR* **2011**, *50*, 209–218.

(505) Didenko, T.; Boelens, R.; Rudiger, S. G. D. 3D DOSY-TROSY to Determine the Translational Diffusion Coefficient of Large Protein Complexes. *Protein Eng. Des. Sel.* **2011**, *24*, 99–103.

(506) Horst, R.; Horwich, A. L.; Wüthrich, K. Translational Diffusion of Macromolecular Assemblies Measured Using Transverse-Relaxation-Optimized Pulsed Field Gradient NMR. *J. Am. Chem. Soc.* **2011**, *133*, 16354–16357.

(507) Lee, J.; Park, S. H.; Cavagnero, S.; Lee, J. H. High-Resolution Diffusion Measurements of Proteins by NMR under Near-Physiological Conditions. *Anal. Chem.* **2020**, *92*, 5073–5081.

(508) Morin, S. A Practical Guide to Protein Dynamics from 15N Spin Relaxation in Solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 245–262.

(509) Jarymowycz, V. A.; Stone, M. J. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. *Chem. Rev.* **2006**, *106*, 1624–1671.

(510) Kay, L. E.; Torchia, D. A.; Bax, A. Backbone Dynamics of Proteins as Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry* **1989**, *28*, 8972–8979.

(511) Jarymowycz, V. A.; Stone, M. J. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. *Chem. Rev.* **2006**, *106*, 1624–1671.

(512) Kneller, J. M.; Lu, M.; Bracken, C. An Effective Method for the Discrimination of Motional Anisotropy and Chemical Exchange. *J. Am. Chem. Soc.* **2002**, *124*, 1852–1853.

(513) Li, C.; Pielak, G. J. Using NMR to Distinguish Viscosity Effects from Nonspecific Protein Binding under Crowded Conditions. *J. Am. Chem. Soc.* **2009**, *131*, 1368–1369.

(514) Wang, Y.; Li, C.; Pielak, G. J. Effects of Proteins on Protein Diffusion. *J. Am. Chem. Soc.* **2010**, *132*, 9392–9397.

(515) Song, X.; An, L.; Wang, M.; Chen, J.; Liu, Z.; Yao, L. Osmolytes Can Destabilize Proteins in Cells by Modulating Electrostatics and Quinary Interactions. *ACS Chem. Biol.* **2021**, *16*, 864–871.

(516) Kim, S.; Wu, K.-P.; Baum, J. Fast Hydrogen Exchange Affects ^{15}N Relaxation Measurements in Intrinsically Disordered Proteins. *J. Biomol. NMR* **2013**, *55*, 249–256.

(517) Tolman, J. R.; Ruan, K. NMR Residual Dipolar Couplings as Probes of Biomolecular Dynamics. *Chem. Rev.* **2006**, *106*, 1720–1736.

(518) Hwang, T. L.; van Zijl, P. C.; Mori, S. Accurate Quantitation of Water-Amide Proton Exchange Rates Using the Phase-Modulated CLEAN Chemical EXchange (CLEANEX-PM) Approach with a Fast-HSQC (FHSQC) Detection Scheme. *J. Biomol. NMR* **1998**, *11*, 221–226.

(519) Modig, K.; Poulsen, F. M. Model-Independent Interpretation of NMR Relaxation Data for Unfolded Proteins: The Acid-Denatured State of ACBP. *J. Biomol. NMR* **2008**, *42*, 163–177.

(520) Adamski, W.; Salvi, N.; Maurin, D.; Magnat, J.; Milles, S.; Jensen, M. R.; Abyzov, A.; Moreau, C. J.; Blackledge, M. A Unified Description of Intrinsically Disordered Protein Dynamics under Physiological Conditions Using NMR Spectroscopy. *J. Am. Chem. Soc.* **2019**, *141*, 17817–17829.

(521) Khan, S. N.; Charlier, C.; Augustyniak, R.; Salvi, N.; Déjean, V.; Bodenhausen, G.; Lequin, O.; Pelupessy, P.; Ferrage, F. Distribution of Pico- and Nanosecond Motions in Disordered Proteins from Nuclear Spin Relaxation. *Biophys. J.* **2015**, *109*, 988–999.

(522) Mandel, A. M.; Akke, M.; Palmer, I., Arthur G. Backbone Dynamics of Escherichia Coli Ribonuclease HI: Correlations with Structure and Function in an Active Enzyme. *J. Mol. Biol.* **1995**, *246*, 144–163.

(523) Palmer, A. G.; Rance, M.; Wright, P. E. Intramolecular Motions of a Zinc Finger DNA-Binding Domain from Xfin Characterized by Proton-Detected Natural Abundance Carbon-13 Heteronuclear NMR Spectroscopy. *J. Am. Chem. Soc.* **1991**, *113*, 4371–4380.

(524) Hansen, D. F.; Vallurupalli, P.; Kay, L. E. Using Relaxation Dispersion NMR Spectroscopy to Determine Structures of Excited, Invisible Protein States. *J. Biomol. NMR* **2008**, *41*, 113–120.

(525) Long, D.; Liu, M.; Yang, D. Accurately Probing Slow Motions on Millisecond Timescales with a Robust NMR Relaxation Experiment. *J. Am. Chem. Soc.* **2008**, *130*, 2432–2433.

(526) Yuwen, T.; Huang, R.; Vallurupalli, P.; Kay, L. E. A Methyl-TROSY-Based 1H Relaxation Dispersion Experiment for Studies of Conformational Exchange in High Molecular Weight Proteins. *Angew Chem Int Ed* **2019**, 5.

(527) Englander, S. W.; Mayne, L.; Kan, Z.-Y.; Hu, W. Protein Folding—How and Why: By Hydrogen Exchange, Fragment Separation, and Mass Spectrometry. *Annu. Rev. Biophys.* **2016**, *45*, 135–152.

(528) Hodge, E. A.; Benhaim, M. A.; Lee, K. K. Bridging Protein Structure, Dynamics, and Function Using Hydrogen/Deuterium-exchange Mass Spectrometry. *Protein Sci.* **2020**, *29*, 843–855.

(529) Dempsey, C. E. Hydrogen Exchange in Peptides and Proteins Using NMR Spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 135–170.

(530) Hernández, G.; LeMaster, D. M. NMR Analysis of Native-State Protein Conformational Flexibility by Hydrogen Exchange. In *Protein Structure, Stability, and Interactions*; Shriver, J. W., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2009; pp 285–310.

(531) Dass, R.; Corlianò, E.; Mulder, F. A. A. The Contribution of Electrostatics to Hydrogen Exchange in the Unfolded Protein State. *Biophys. J.* **2021**, S0006349521006482.

(532) James, E. I.; Murphree, T. A.; Vorauer, C.; Engen, J. R.; Guttman, M. Advances in Hydrogen/Deuterium Exchange Mass Spectrometry and the Pursuit of Challenging Biological Systems. *Chem. Rev.* **2021**, acs.chemrev.1c00279.

(533) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins Struct. Funct. Bioinforma.* **1993**, *17*, 75–86.

(534) Chevelkov, V.; Xue, Y.; Krishna Rao, D.; Forman-Kay, J. D.; Skrynnikov, N. R. 15NH/D-SOLEXSY Experiment for Accurate Measurement of Amide Solvent Exchange Rates: Application to Denatured DrkN SH3. *J. Biomol. NMR* **2010**, *46*, 227–244.

(535) Rennella, E.; Solyom, Z.; Brutscher, B. Measuring Hydrogen Exchange in Proteins by Selective Water Saturation in 1H–15N SOFAST/BEST-Type Experiments: Advantages and Limitations. *J. Biomol. NMR* **2014**, *60*, 99–107.

(536) Dass, R.; Corlianò, E.; Mulder, F. A. A. Measurement of Very Fast Exchange Rates of Individual Amide Protons in Proteins by NMR Spectroscopy. *ChemPhysChem* **2019**, *20*, 231–235.

(537) Gossert, A. D.; Jahnke, W. NMR in Drug Discovery: A Practical Guide to Identification and Validation of Ligands Interacting with Biological Macromolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *97*, 82–125.

(538) Hanzawa, H.; Shimada, T.; Takahashi, M.; Takahashi, H. Revisiting Biomolecular NMR Spectroscopy for Promoting Small-Molecule Drug Discovery. *J. Biomol. NMR* **2020**, *138*, 4539–8.

(539) Dalvit, C.; Veronesi, M.; Vulpetti, A. Fluorine NMR Functional Screening: From Purified Enzymes to Human Intact Living Cells. *J. Biomol. NMR* **2020**, *74*, 613–631.

(540) Stott, K.; Stonehouse, J.; Keeler, J.; Hwang, T.-L.; Shaka, A. J. Excitation Sculpting in High-Resolution Nuclear Magnetic Resonance Spectroscopy: Application to Selective NOE Experiments. *J. Am. Chem. Soc.* **1995**, *117*, 4199–4200.

(541) Ma, J.; McLeod, S.; MacCormack, K.; Sriram, S.; Gao, N.; Breeze, A. L.; Hu, J. Real-Time Monitoring of New Delhi Metallo-β-Lactamase Activity in Living Bacterial Cells by 1H NMR Spectroscopy. *Angew. Chem. Int. Ed Engl.* **2014**, *53*, 2130–2133.

(542) Ma, J.; Cao, Q.; McLeod, S. M.; Ferguson, K.; Gao, N.; Breeze, A. L.; Hu, J. Target-Based Whole-Cell Screening by 1H NMR Spectroscopy. *Angew. Chem.* **2015**, *127*, 4846–4849.

(543) Veronesi, M.; Giacomina, F.; Romeo, E.; Castellani, B.; Ottonello, G.; Lambruschini, C.; Garau, G.; Scarpelli, R.; Bandiera, T.; Piomelli, D.; et al. Fluorine Nuclear Magnetic Resonance-Based Assay in Living Mammalian Cells. *Anal. Biochem.* **2016**, *495*, 52–59.

(544) Bhunia, A.; Bhattacharjya, S.; Chatterjee, S. Applications of Saturation Transfer Difference NMR in Biological Systems. *Drug Discov. Today* **2012**, *17*, 505–513.

(545) Wagstaff, J. L.; Taylor, S. L.; Howard, M. J. Recent Developments and Applications of Saturation Transfer Difference Nuclear Magnetic Resonance (STD NMR) Spectroscopy. *Mol. Biosyst.* **2013**, *9*, 571–577.

(546) Meyer, B.; Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angew. Chem.* **2003**, *42*, 864–890.

(547) Bouvier, G.; Simenel, C.; Jang, J.; Kalia, N. P.; Choi, I.; Nilges, M.; Pethe, K.; Izadi-Pruneyre, N. Target Engagement and Binding Mode of an Antituberculosis Drug to Its Bacterial Target Deciphered in Whole Living Cells by NMR. *Biochemistry* **2019**, *58*, 526–533.

(548) Rademacher, C.; Krishna, N. R.; Palcic, M.; Parra, F.; Peters, T. NMR Experiments Reveal the Molecular Basis of Receptor Recognition by a Calicivirus. *J. Am. Chem. Soc.* **2008**, *130*, 3669–3675.

(549) Rademacher, C.; Peters, T. Molecular Recognition of Ligands by Native Viruses and Virus-like Particles as Studied by NMR Experiments. *Top. Curr. Chem.* **2008**, *273*, 183–202.

(550) Cutting, B.; Shelke, S. V.; Dragic, Z.; Wagner, B.; Gathje, H.; Kelm, S.; Ernst, B. Sensitivity Enhancement in Satura-

tion Transfer Difference (STD) Experiments through Optimized Excitation Schemes. *Magn. Reson. Chem.* **2007**, *45*, 720–724.

(551) Claasen, B.; Axmann, M.; Meinecke, R.; Meyer, B. Direct Observation of Ligand Binding to Membrane Proteins in Living Cells by a Saturation Transfer Double Difference (STDD) NMR Spectroscopy Method Shows a Significantly Higher Affinity of Integrin Alpha(IIb)Beta3 in Native Platelets than in Liposomes. *J. Am. Chem. Soc.* **2005**, *127*, 916–919.

(552) Pereira, A.; Pfeifer, T. A.; Grigliatti, T. A.; Andersen, R. J. Functional Cell-Based Screening and Saturation Transfer Double-Difference NMR Have Identified Haplosamate A as a Cannabinoid Receptor Agonist. *ACS Chem. Biol.* **2009**, *4*, 139–144.

(553) Airoldi, C.; Giovannardi, S.; La Ferla, B.; Jiménez-Barbero, J.; Nicotra, F. Saturation Transfer Difference NMR Experiments of Membrane Proteins in Living Cells under HR-MAS Conditions: The Interaction of the SGLT1 Co-Transporter with Its Ligands. *Chem. - Eur. J.* **2011**, *17*, 13395–13399.

(554) Li, S.; Ahmed, L.; Zhang, R.; Pan, Y.; Matsunami, H.; Burger, J. L.; Block, E.; Batista, V. S.; Zhuang, H. Smelling Sulfur: Copper and Silver Regulate the Response of Human Odorant Receptor OR2T11 to Low-Molecular-Weight Thiols. *J. Am. Chem. Soc.* **2016**, *138*, 13281–13288.

(555) Mari, S.; Serrano-Gómez, D.; Cañada, F. J.; Corbí, A. L.; Jiménez-Barbero, J. 1D Saturation Transfer Difference NMR Experiments on Living Cells: The DC-SIGN/Oligomannose Interaction. *Angew. Chem.* **2005**, *44*, 296–298.

(556) Burger, J. L.; Jeerage, K. M.; Bruno, T. J. Direct Nuclear Magnetic Resonance Observation of Odorant Binding to Mouse Odorant Receptor MOR244-3. *Anal. Biochem.* **2016**, *502*, 64–72.

(557) Vasile, F.; Gubinelli, F.; Panigada, M.; Soprana, E.; Siccardi, A.; Potenza, D. NMR Interaction Studies of Neu5Ac-α-(2,6)-Gal-β-(1-4)-GlcNAc with Influenza-Virus Hemagglutinin Expressed in Transfected Human Cells. *Glycobiology* **2018**, *28*, 42–49.

(558) Madge, P. D.; Maggioni, A.; Pascolutti, M.; Amin, M.; Waespy, M.; Bellette, B.; Thomson, R. J.; Kelm, S.; von Itzstein, M.; Haselhorst, T. Structural Characterisation of High Affinity Siglec-2 (CD22) Ligands in Complex with Whole Burkitt's Lymphoma (BL) Daudi Cells by NMR Spectroscopy. *Sci. Rep.* **2016**, *6*, 36012–12.

(559) Jayalakshmi, V.; Krishna, N. R. Complete Relaxation and Conformational Exchange Matrix (CORCEMA) Analysis of Intermolecular Saturation Transfer Effects in Reversibly Forming Ligand-Receptor Complexes. *J. Magn. Reson.* **2002**, *155*, 106–118.

(560) Jayalakshmi, V.; Rama Krishna, N. CORCEMA Refinement of the Bound Ligand Conformation within the Protein Binding Pocket in Reversibly Forming Weak Complexes Using STD-NMR Intensities. *J. Magn. Reson.* **2004**, *168*, 36–45.

(561) Potenza, D.; Vasile, F.; Belvisi, L.; Civera, M.; Araldi, E. M. V. STD and TrNOESY NMR Study of Receptor-Ligand Interactions in Living Cancer Cells. *ChemBioChem* **2011**, *12*, 695–699.

(562) Mari, S.; Invernizzi, C.; Spitaleri, A.; Alberici, L.; Ghitti, M.; Bordignon, C.; Traversari, C.; Rizzardi, G.-P.; Musco, G. 2D TR-NOESY Experiments Interrogate and Rank Ligand-Receptor Interactions in Living Human Cancer Cells. *Angew. Chem.* **2010**, *49*, 1071–1074.

(563) Brancaccio, D.; Diana, D.; Di Maro, S.; Di Leva, F. S.; Tomassi, S.; Fattorusso, R.; Russo, L.; Scala, S.; Trotta, A. M.; Portella, L.; et al. Ligand-Based NMR Study of C-X-C Chemokine Receptor Type 4 (CXCR4)-Ligand Interactions on Living Cancer Cells. *J. Med. Chem.* **2018**, *61*, 2910–2923.

(564) Di Stasi, R.; Diana, D.; Capasso, D.; Di Gaetano, S.; De Rosa, L.; Celentano, V.; Isernia, C.; Fattorusso, R.; D'Andrea, L. D. VEGFR Recognition Interface of a Proangiogenic VEGF-Mimetic Peptide Determined In Vitro and in the Presence of Endothelial Cells by NMR Spectroscopy. *Chem. - Eur. J.* **2018**, *24*, 11461–11466.

(565) Primikyri, A.; Sayyad, N.; Quilici, G.; Vrettos, E. I.; Lim, K.; Chi, S. W.; Musco, G.; Gerothanassis, I. P.; Tzakos, A. G. Probing the Interaction of a Quercetin Bioconjugate with Bcl-2 in Living Human Cancer Cells with in-Cell NMR Spectroscopy. *FEBS Lett.* **2018**, *592*, 3367–3379.

(566) Antanasijevic, A.; Kingsley, C.; Basu, A.; Bowlin, T. L.; Rong, L.; Caffrey, M. Application of Virus-like Particles (VLP) to NMR Characterization of Viral Membrane Protein Interactions. *J. Biomol. NMR* **2016**, *64*, 255–265.

(567) Palmioli, A.; Ceresa, C.; Tripodi, F.; La Ferla, B.; Nicolini, G.; Airoldi, C. On-Cell Saturation Transfer Difference NMR Study of Bombesin Binding to GRP Receptor. *Bioorganic Chem.* **2020**, 103861.

(568) Guzzetti, I.; Civera, M.; Vasile, F.; Araldi, E. M.; Belvisi, L.; Gennari, C.; Potenza, D.; Fanelli, R.; Piarulli, U. Determination of the Binding Epitope of RGD-Peptidomimetics to Avβ3 and α(IIb)B3 Integrin-Rich Intact Cells by NMR and Computational Studies. *Org. Biomol. Chem.* **2013**, *11*, 3886–3893.

(569) Vasile, F.; Menchi, G.; Lenci, E.; Guarna, A.; Potenza, D.; Trabocchi, A. Insight to the Binding Mode of Triazole RGD-Peptidomimetics to Integrin-Rich Cancer Cells by NMR and Molecular Modeling. *Bioorg. Med. Chem.* **2016**, *24*, 989–994.

(570) Guzzetti, I.; Civera, M.; Vasile, F.; Arosio, D.; Tringali, C.; Piarulli, U.; Gennari, C.; Pignataro, L.; Belvisi, L.; Potenza, D. Insights into the Binding of Cyclic RGD Peptidomimetics to A5β1 Integrin by Using Live-Cell NMR And Computational Studies. *ChemistryOpen* **2017**, *6*, 128–136.

(571) Clore, G. M.; Gronenborn, A. M. Theory and Applications of the Transferred Nuclear Overhauser Effect to the Study of the Conformations of Small Ligands Bound to Proteins. *J. Magn. Reson. 1969* **1982**, *48*, 402–417.

(572) Orts, J.; Gossert, A. D. Structure Determination of Protein-Ligand Complexes by NMR in Solution. *Methods* **2018**, *138–139*, 3–25.

(573) Valentinis, B.; Porcellini, S.; Asperti, C.; Cota, M.; Zhou, D.; Di Matteo, P.; Garau, G.; Zucchelli, C.; Avanzi, N. R.; Rizzardi, G.-P.; et al. Mechanism of Action of the Tumor Vessel Targeting Agent NGR-HTNF: Role of Both NGR Peptide and HTNF in Cell Binding and Signaling. *Int. J. Mol. Sci.* **2019**, *20*.

(574) Zheng, G.; Price, W. S. Solvent Signal Suppression in NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2010**, *56*, 267–288.

(575) Lam, B.; Simpson, A. J. Direct 1H NMR Spectroscopy of Dissolved Organic Matter in Natural Waters. *Analyst* **2008**, *133*, 263–269.

(576) Soong, R.; Nagato, E.; Sutrisno, A.; Fortier-McGill, B.; Akhter, M.; Schmidt, S.; Heumann, H.; Simpson, A. J. In Vivo NMR Spectroscopy: Toward Real Time Monitoring of Environmental Stress. *Magn. Reson. Chem.* **2015**, *53*, 774–779.

(577) Aguilar, J. A.; Kenwright, Simon. J. Robust NMR Water Signal Suppression for Demanding Analytical Applications. *Analyst* **2016**, *141*, 236–242.

(578) Aguilar, J. A.; Cassani, J.; Probert, F.; Palace, J.; Claridge, T. D. W.; Botana, A.; Kenwright, A. M. Reliable, High-Quality Suppression of NMR Signals Arising from Water and Macromolecules: Application to Bio-Fluid Analysis. *Analyst* **2019**, *144*, 7270–7277.

(579) Le Guennec, A.; Tayyari, F.; Edison, A. S. Alternatives to Nuclear Overhauser Enhancement Spectroscopy Presat and Carr–Purcell–Meiboom–Gill Presat for NMR-Based Metabolomics. *Anal. Chem.* **2017**, *89*, 8582–8588.

(580) Mayer, M.; Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR to Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.* **2001**, *123*, 6108–6117.

(581) Scherf, T.; Anglister, J. A T1 Rho-Filtered Two-Dimensional Transferred NOE Spectrum for Studying Antibody Interactions with Peptide Antigens. *Biophys. J.* **1993**, *64*, 754–761.

(582) Piersanti, E.; Rezig, L.; Tranchida, F.; El-Houri, W.; Abagana, S. M.; Campredon, M.; Shintu, L.; Yemloul, M. Evaluation of the Rotating-Frame Relaxation ($T_{1\rho}$) Filter and Its Application in Metabolomics as an Alternative to the Transverse Relaxation (T_2) Filter. *Anal. Chem.* **2021**, *93*, 8746–8753.

(583) Garcia-Perez, I.; Posma, J. M.; Serrano-Contreras, J. I.; Boulangé, C. L.; Chan, Q.; Frost, G.; Stamler, J.; Elliott, P.; Lindon, J. C.; Holmes, E.; et al. Identifying Unknown Metabolites Using

NMR-Based Metabolic Profiling Techniques. *Nat. Protoc.* **2020**, *15*, 2538–2567.

(584) Tang, H.; Wang, Y.; Nicholson, J. K.; Lindon, J. C. Use of Relaxation-Edited One-Dimensional and Two Dimensional Nuclear Magnetic Resonance Spectroscopy to Improve Detection of Small Metabolites in Blood Plasma. *Anal. Biochem.* **2004**, *325*, 260–272.

(585) Aguilar, J. A.; Nilsson, M.; Bodenhausen, G.; Morris, G. A. Spin Echo NMR Spectra without J Modulation. *Chem. Commun.* **2012**, *48*, 811–813.

(586) Hassan, Q.; Dutta Majumdar, R.; Wu, B.; Lane, D.; Tabatabaei-Anraki, M.; Soong, R.; Simpson, M. J.; Simpson, A. J. Improvements in Lipid Suppression for 1H NMR-Based Metabolomics: Applications to Solution-State and HR-MAS NMR in Natural and in Vivo Samples. *Magn. Reson. Chem.* **2019**, *57*, 69–81.

(587) Li, T.; Liao, Y.; Jiang, X.; Mu, D.; Hou, X.; Zhang, C.; Deng, P. PH Detection in Biological Samples by 1D and 2D 1H–31P NMR. *Talanta* **2018**, *178*, 538–544.

(588) Jin, X.; Kang, S.; Tanaka, S.; Park, S. Monitoring the Glutathione Redox Reaction in Living Human Cells by Combining Metabolic Labeling with Heteronuclear NMR. *Angew. Chem. Int. Ed Engl.* **2016**, *55*, 7939–7942.

(589) Cox, N.; Kuemmerle, R.; Millard, P.; Cahoreau, E.; François, J.-M.; Parrou, J.-L.; Lippens, G. Integrated PH Measurement during Reaction Monitoring with Dual-Reception 1H-31P NMR Spectroscopy. *Anal. Chem.* **2019**, *91*, 3959–3963.

(590) Nikolaev, Y.; Ripin, N.; Soste, M.; Picotti, P.; Iber, D.; Allain, F. H.-T. Systems NMR: Single-Sample Quantification of RNA, Proteins and Metabolites for Biomolecular Network Analysis. *Nat. Methods* **2019**, *16*, 743–749.

(591) Guin, D.; Gruebele, M. Weak Chemical Interactions That Drive Protein Evolution: Crowding, Sticking, and Quinary Structure in Folding and Function. *Chem. Rev.* **2019**, *119*, 10691–10717.

(592) Sridharan, S.; Kurzawa, N.; Werner, T.; Günthner, I.; Helm, D.; Huber, W.; Bantscheff, M.; Savitski, M. M. Proteome-Wide Solubility and Thermal Stability Profiling Reveals Distinct Regulatory Roles for ATP. *Nat. Commun.* **2019**, 1–13.

(593) Tokuriki, N.; Tawfik, D. S. Stability Effects of Mutations and Protein Evolvability. *Curr. Opin. Struct. Biol.* **2009**, *19*, 596–604.

(594) Goldenzweig, A.; Fleishman, S. J. Principles of Protein Stability and Their Application in Computational Design. *Annu. Rev. Biochem.* **2018**, *87*, 105–129.

(595) Nisthal, A.; Wang, C. Y.; Ary, M. L.; Mayo, S. L. Protein Stability Engineering Insights Revealed by Domain-Wide Comprehensive Mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 16367–16377.

(596) Xavier, J. S.; Nguyen, T.-B.; Karmarkar, M.; Portelli, S.; Rezende, P. M.; Velloso, J. P. L.; Ascher, D. B.; Pires, D. E. V. ThermoMutDB: A Thermodynamic Database for Missense Mutations. *Nucleic Acids Res.* **2021**, *49*, D475–D479.

(597) Zhang, S.; Greening, D. W.; Hong, Y. Recent Advances in Bioanalytical Methods to Measure Proteome Stability in Cells. *Analyst* **2021**, *146*, 2097–2109.

(598) Caldararu, O.; Blundell, T. L.; Kepp, K. P. A Base Measure of Precision for Protein Stability Predictors: Structural Sensitivity. *BMC Bioinformatics* **2021**, 1–14.

(599) Marabotti, A.; Del Prete, E.; Scafuri, B.; Facchiano, A. Performance of Web Tools for Predicting Changes in Protein Stability Caused by Mutations. *BMC Bioinformatics* **2021**, 1–19.

(600) Broom, A.; Trainor, K.; Jacobi, Z.; Meiering, E. M. Computational Modeling of Protein Stability: Quantitative Analysis Reveals Solutions to Pervasive Problems. *Struct. Lond. Engl. 1993* **2020**, *28*, 717-726.e3.

(601) Powers, E. T.; Gierasch, L. M. The Proteome Folding Problem and Cellular Proteostasis. *J. Mol. Biol.* **2021**, *433*, 167197.

(602) Leuenberger, P.; Ganscha, S.; Kahraman, A.; Cappelletti, V.; Boersema, P. J.; von Mering, C.; Claassen, M.; Picotti, P. Cell-Wide Analysis of Protein Thermal Unfolding Reveals Determinants of Thermostability. *Science* **2017**, *355*.

(603) Jarzab, A.; Kurzawa, N.; Hopf, T.; Moerch, M.; Zecha, J.; Leijten, N.; Bian, Y.; Musiol, E.; Maschberger, M.; Stoehr, G.; et al. Meltome Atlas—Thermal Proteome Stability across the Tree of Life. *Nat. Methods* **2020**, 1–15.

(604) Pal, L. R.; Moult, J. Genetic Basis of Common Human Disease: Insight into the Role of Missense SNPs from Genome-Wide Association Studies. *J. Mol. Biol.* **2015**, *427*, 2271–2289.

(605) Redler, R. L.; Das, J.; Diaz, J. R.; Dokholyan, N. V. Protein Destabilization as a Common Factor in Diverse Inherited Disorders. *J. Mol. Evol.* **2016**, *82*, 11–16.

(606) Stein, A.; Fowler, D. M.; Hartmann-Petersen, R.; Lindorff-Larsen, K. Biophysical and Mechanistic Models for Disease-Causing Protein Variants. *Trends Biochem. Sci.* **2019**, *44*, 575–588.

(607) Sanavia, T.; Birolo, G.; Montanucci, L.; Turina, P.; Capriotti, E.; Fariselli, P. Limitations and Challenges in Protein Stability Prediction upon Genome Variations: Towards Future Applications in Precision Medicine. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1968–1979.

(608) Myers, J. K.; Nick Pace, C.; Martin Scholtz, J. Denaturant m Values and Heat Capacity Changes: Relation to Changes in Accessible Surface Areas of Protein Unfolding. *Protein Sci.* **1995**, *4*, 2138–2148.

(609) Prabhu, N. V.; Sharp, K. A. HEAT CAPACITY IN PROTEINS. *Annu. Rev. Phys. Chem.* **2005**, *56*, 521–548.

(610) Senske, M.; Törk, L.; Born, B.; Havenith, M.; Herrmann, C.; Ebbinghaus, S. Protein Stabilization by Macromolecular Crowding through Enthalpy Rather Than Entropy. *J. Am. Chem. Soc.* **2014**, *136*, 9036–9041.

(611) Katava, M.; Stirnemann, G.; Pachetti, M.; Capaccioli, S.; Paciaroni, A.; Sterpone, F. Specific Interactions and Environment Flexibility Tune Protein Stability under Extreme Crowding. *J. Phys. Chem. B* **2021**, *125*, 6103–6111.

(612) Stadmiller, S. S.; Pielak, G. J. Enthalpic Stabilization of an SH3 Domain by D_2O. *Protein Sci.* **2018**, *27*, 1710–1716.

(613) Tabaka, M.; Kalwarczyk, T.; Szymanski, J.; Hou, S.; Holyst, R. The Effect of Macromolecular Crowding on Mobility of Biomolecules, Association Kinetics, and Gene Expression in Living Cells. *Front. Phys.* **2014**, *2*.

(614) Wachsmuth, M.; Conrad, C.; Bulkescher, J.; Koch, B.; Mahen, R.; Isokane, M.; Pepperkok, R.; Ellenberg, J. High-Throughput Fluorescence Correlation Spectroscopy Enables Analysis of Proteome Dynamics in Living Cells. *Nat. Biotechnol.* **2015**, *33*, 384–389.

(615) Bubak, G.; Kwapiszewska, K.; Kalwarczyk, T.; Bielec, K.; Andryszewski, T.; Iwan, M.; Bubak, S.; Hołyst, R. Quantifying Nanoscale Viscosity and Structures of Living Cells Nucleus from Mobility Measurements. *J. Phys. Chem. Lett.* **2021**, *12*, 294–301.

(616) Xiang, L.; Chen, K.; Yan, R.; Li, W.; Xu, K. Single-Molecule Displacement Mapping Unveils Nanoscale Heterogeneities in Intracellular Diffusivity. *Nat. Methods* **2020**, *17*, 524–530.

(617) Umuhire Juru, A.; Patwardhan, N. N.; Hargrove, A. E. Understanding the Contributions of Conformational Changes, Thermodynamics, and Kinetics of RNA–Small Molecule Interactions. *ACS Chem. Biol.* **2019**, *14*, 824–838.

(618) Marušič, M.; Schlagnitweit, J.; Petzold, K. RNA Dynamics by NMR Spectroscopy. *ChemBioChem* **2019**, *20*, 2685–2710.

(619) Liu, B.; Shi, H.; Al-Hashimi, H. M. Developments in Solution-State NMR Yield Broader and Deeper Views of the Dynamic Ensembles of Nucleic Acids. *Curr. Opin. Struct. Biol.* **2021**, *70*, 16–25.

(620) Takahashi, S.; Yamamoto, J.; Kitamura, A.; Kinjo, M.; Sugimoto, N. Characterization of Intracellular Crowding Environments with Topology-Based DNA Quadruplex Sensors. *Anal. Chem.* **2019**, *91*, 2586–2590.

(621) Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The Regulation and Functions of DNA and RNA G-Quadruplexes. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 459–474.

(622) Jana, J.; Weisz, K. Thermodynamic Stability of G-Quadruplexes: Impact of Sequence and Environment. *ChemBioChem* **2021**, *22*, 2848–2856.

(623) Takahashi, S.; Sugimoto, N. Stability Prediction of Canonical and Non-Canonical Structures of Nucleic Acids in Various Molecular Environments and Cells. *Chem. Soc. Rev.* **2020**, *49*, 8439–8468.

(624) Nakano, S.; Miyoshi, D.; Sugimoto, N. Effects of Molecular Crowding on the Structures, Interactions, and Functions of Nucleic Acids. *Chem. Rev.* **2014**, *114*, 2733–2758.

(625) Morag, O.; Abramov, G.; Goldbourt, A. Complete Chemical Shift Assignment of the SsDNA in the Filamentous Bacteriophage Fd Reports on Its Conformation and on Its Interface with the Capsid Shell. *J. Am. Chem. Soc.* **2014**, *136*, 2292–2301.

(626) Goldbourt, A. Structural Characterization of Bacteriophage Viruses by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *114–115*, 192–210.

(627) Quinn, C. M.; Lu, M.; Suiter, C. L.; Hou, G.; Zhang, H.; Polenova, T. Magic Angle Spinning NMR of Viruses. *Prog. Nucl. Magn. Reson. Spectrosc.* **2015**, *86–87*, 21–40.

(628) Lu, M.; Wang, M.; Sergeyev, I. V.; Quinn, C. M.; Struppe, J.; Rosay, M.; Maas, W.; Gronenborn, A. M.; Polenova, T. 19F Dynamic Nuclear Polarization at Fast Magic Angle Spinning for NMR of HIV-1 Capsid Protein Assemblies. *J. Am. Chem. Soc.* **2019**, *141*, 5681–5691.

(629) Hänsel, R.; Löhr, F.; Trantirek, L.; Dötsch, V. High-Resolution Insight into G-Overhang Architecture. *J. Am. Chem. Soc.* **2013**, *135*, 2816–2824.

(630) Ishizuka, T.; Zhao, P.-Y.; Bao, H.-L.; Xu, Y. A Multi-Functional Guanine Derivative for Studying the DNA G-Quadruplex Structure. *Analyst* **2017**, *142*, 4083–4088.

(631) 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.

(632) Yamaoki, Y.; Nagata, T.; Sakamoto, T.; Katahira, M. Recent Progress of In-Cell NMR of Nucleic Acids in Living Human Cells. *Biophys. Rev.* **2020**, *12*, 411–417.

(633) Giassa, I.-C.; Rynes, J.; Fessl, T.; Foldynova-Trantirkova, S.; Trantirek, L. Advances in the Cellular Structural Biology of Nucleic Acids. *FEBS Lett.* **2018**, *592*, 1997–2011.

(634) Cheng, M.; Qiu, D.; Tamon, L.; Ištvánková, E.; Víšková, P.; Amrane, S.; Guédin, A.; Chen, J.; Lacroix, L.; Ju, H.; et al. Thermal and PH Stabilities of i-DNA: Confronting in Vitro Experiments with Models and In-Cell NMR Data. *Angew. Chem. Int. Ed.* **2021**, *60*, 10286–10294.

(635) Krafčík, D.; Ištvánková, E.; Džatko, Š.; Víšková, P.; Foldynová-Trantírková, S.; Trantírek, L. Towards Profiling of the G-Quadruplex Targeting Drugs in the Living Human Cells Using NMR Spectroscopy. *Int. J. Mol. Sci.* **2021**, *22*, 6042.

(636) Reddy Sannapureddi, R. K.; Mohanty, M. K.; Gautam, A. K.; Sathyamoorthy, B. Characterization of DNA G-Quadruplex Topologies with NMR Chemical Shifts. *J. Phys. Chem. Lett.* **2020**, *11*, 10016–10022.

(637) Deprey, K.; Batistatou, N.; Kritzer, J. A. A Critical Analysis of Methods Used to Investigate the Cellular Uptake and Subcellular Localization of RNA Therapeutics. *Nucleic Acids Res.* **2020**, *48*, 7623–7639.

(638) Hatos, A.; Hajdu-Soltész, B.; Monzon, A. M.; Palopoli, N.; Álvarez, L.; Aykac-Fas, B.; Bassot, C.; Benítez, G. I.; Bevilacqua, M.; Chasapi, A.; et al. DisProt: Intrinsic Protein Disorder Annotation in 2020. *Nucleic Acids Res.* **2019**, gkz975.

(639) Krystkowiak, I.; Davey, N. E. SLiMSearch: A Framework for Proteome-Wide Discovery and Annotation of Functional Modules in Intrinsically Disordered Regions. *Nucleic Acids Res.* **2017**, *45*, W464–W469.

(640) Davey, N. E. ScienceDirect The Functional Importance of Structure in Unstructured Protein Regions. *Curr. Opin. Struct. Biol.* **2019**, *56*, 155–163.

(641) Teilum, K.; Olsen, J. G.; Kragelund, B. B. On the Specificity of Protein–Protein Interactions in the Context of Disorder. *Biochem. J.* **2021**, *478*, 2035–2050.

(642) Macossay-Castillo, M.; Marvelli, G.; Guharoy, M.; Jain, A.; Kihara, D.; Tompa, P.; Wodak, S. J. The Balancing Act of Intrinsically Disordered Proteins: Enabling Functional Diversity While Minimizing Promiscuity. *J. Mol. Biol.* **2019**, *431*, 1650–1670.

(643) Babu, M. M. The Contribution of Intrinsically Disordered Regions to Protein Function, Cellular Complexity, and Human Disease. *Biochem. Soc. Trans.* **2016**, *44*, 1185–1200.

(644) Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, *176*, 419–434.

(645) Boija, A.; Klein, I. A.; Sabari, B. R.; Dall'Agnese, A.; Coffey, E. L.; Zamudio, A. V.; Li, C. H.; Shrinivas, K.; Manteiga, J. C.; Hannett, N. M.; et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **2018**, 1–31.

(646) Sørensen, C. S.; Jendroszek, A.; Kjaergaard, M. Linker Dependence of Avidity in Multivalent Interactions Between Disordered Proteins. *J. Mol. Biol.* **2019**, *431*, 4784–4795.

(647) Huang, Q.; Li, M.; Lai, L.; Liu, Z. Allostery of Multidomain Proteins with Disordered Linkers. *Curr. Opin. Struct. Biol.* **2020**, *62*, 175–182.

(648) Zarin, T.; Strome, B.; Peng, G.; Pritišanac, I.; Forman-Kay, J. D.; Moses, A. M. Identifying Molecular Features That Are Associated with Biological Function of Intrinsically Disordered Protein Regions. *eLife* **2021**, *10*, e60220.

(649) Uversky, V. N. Recent Developments in the Field of Intrinsically Disordered Proteins: Intrinsic Disorder–Based Emergence in Cellular Biology in Light of the Physiological and Pathological Liquid–Liquid Phase Transitions. *Annu. Rev. Biophys.* **2021**, *50*, 135–156.

(650) Malinovska, L.; Kroschwald, S.; Alberti, S. Protein Disorder, Prion Propensities, and Self-Organizing Macromolecular Collectives. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2013**, *1834*, 918–931.

(651) Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M. The Amyloid State and Its Association with Protein Misfolding Diseases. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 384–396.

(652) Dyson, H. J.; Wright, P. E. NMR Illuminates Intrinsic Disorder. *Curr. Opin. Struct. Biol.* **2021**, *70*, 44–52.

(653) Schneider, R.; Blackledge, M.; Jensen, M. R. Elucidating Binding Mechanisms and Dynamics of Intrinsically Disordered Protein Complexes Using NMR Spectroscopy. *Curr. Opin. Struct. Biol.* **2019**, *54*, 10–18.

(654) Wang, W.; Perovic, I.; Chittuluru, J.; Kaganovich, A.; Nguyen, L. T. T.; Liao, J.; Auclair, J. R.; Johnson, D.; Landeru, A.; Simorellis, A. K.; et al. A Soluble -Synuclein Construct Forms a Dynamic Tetramer. *Proc. Natl. Acad. Sci.* **2011**, *108*, 17797–17802.

(655) Smith, M. J.; Marshall, C. B.; Theillet, F.-X.; Binolfi, A.; Selenko, P.; Ikura, M. Real-Time NMR Monitoring of Biological Activities in Complex Physiological Environments. *Curr. Opin. Struct. Biol.* **2015**, *32*, 39–47.

(656) Prabakaran, S.; Everley, R. A.; Landrieu, I.; Wieruszeski, J.-M.; Lippens, G.; Steen, H.; Gunawardena, J. Comparative Analysis of Erk Phosphorylation Suggests a Mixed Strategy for Measuring Phospho-Form Distributions. *Mol. Syst. Biol.* **2011**, *7*, 1–15.

(657) Cordier, F.; Chaffotte, A.; Terrien, E.; Prehaud, C.; Theillet, F.-X.; Delepierre, M.; Lafon, M.; Buc, H.; Wolff, N. Ordered Phosphorylation Events in Two Independent Cascades of the PTEN C-Tail Revealed by NMR. *J. Am. Chem. Soc.* **2012**, *134*, 20533–20543.

(658) Theillet, F.-X.; Rose, H. M.; Liokatis, S.; Binolfi, A.; Thongwichian, R.; Stuiver, M.; Selenko, P. Site-Specific NMR Mapping and Time-Resolved Monitoring of Serine and Threonine Phosphorylation in Reconstituted Kinase Reactions and Mammalian Cell Extracts. *Nat. Protoc.* **2013**, *8*, 1416–1432.

(659) Mylona, A.; Theillet, F.-X.; Foster, C.; Cheng, T. M.; Miralles, F.; Bates, P. A.; Selenko, P.; Treisman, R. Opposing Effects of Elk-1 Multisite Phosphorylation Shape Its Response to ERK Activation. *Science* **2016**, *354*, 233–237.

(660) Julien, M.; Bouguechtouli, C.; Alik, A.; Ghouil, R.; Zinn-Justin, S.; Theillet, F.-X. Multiple Site-Specific Phosphorylation of

IDPs Monitored by NMR. In *Intrinsically Disordered Proteins: Methods and Protocols*; Kragelund, B. B., Skriver, K., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2020; pp 793–817.

(661) Dose, A.; Liokatis, S.; Theillet, F.-X.; Selenko, P.; Schwarzer, D. NMR Profiling of Histone Deacetylase and Acetyl-Transferase Activities in Real Time. *ACS Chem. Biol.* **2011**, *6*, 419–424.

(662) Liokatis, S.; Stützer, A.; Elsässer, S. J.; Theillet, F.-X.; Klingberg, R.; van Rossum, B.; Schwarzer, D.; Allis, C. D.; Fischle, W.; Selenko, P. Phosphorylation of Histone H3 Ser10 Establishes a Hierarchy for Subsequent Intramolecular Modification Events. *Nat. Struct. Mol. Biol.* **2012**, *19*, 819–823.

(663) Kamah, A.; Huvent, I.; Cantrelle, F.-X.; Qi, H.; Lippens, G.; Landrieu, I.; Smet-Nocca, C. Nuclear Magnetic Resonance Analysis of the Acetylation Pattern of the Neuronal Tau Protein. *Biochemistry* **2014**, *53*, 3020–3032.

(664) Theillet, F.-X.; Liokatis, S.; Jost, J. O.; Bekei, B.; Rose, H. M.; Binolfi, A.; Schwarzer, D.; Selenko, P. Site-Specific Mapping and Time-Resolved Monitoring of Lysine Methylation by High-Resolution NMR Spectroscopy. *J. Am. Chem. Soc.* **2012**, *134*, 7616–7619.

(665) Altincekic, N.; Lohr, F.; Meier-Credo, J.; Langer, J. D.; Hengesbach, M.; Richter, C.; Schwalbe, H. Site-Specific Detection of Arginine Methylation in Highly Repetitive Protein Motifs of Low Sequence Complexity by NMR. *J. Am. Chem. Soc.* **2020**, *142*, 7647–7654.

(666) Hinterholzer, A.; Stanojlovic, V.; Regl, C.; Huber, C. G.; Cabrele, C.; Schubert, M. Identification and Quantification of Oxidation Products in Full-Length Biotherapeutic Antibodies by NMR Spectroscopy. *Anal. Chem.* **2020**, *92*, 9666–9673.

(667) Grassi, L.; Regl, C.; Wildner, S.; Gadermaier, G.; Huber, C. G.; Cabrele, C.; Schubert, M. Complete NMR Assignment of Succinimide and Its Detection and Quantification in Peptides and Intact Proteins. *Anal. Chem.* **2017**, *89*, 11962–11970.

(668) Hinterholzer, A.; Stanojlovic, V.; Cabrele, C.; Schubert, M. Unambiguous Identification of Pyroglutamate in Full-Length Biopharmaceutical Monoclonal Antibodies by NMR Spectroscopy. *Anal. Chem.* **2019**, *91*, 14299–14305.

(669) Hinterholzer, A.; Stanojlovic, V.; Regl, C.; Huber, C. G.; Cabrele, C.; Schubert, M. Detecting Aspartate Isomerization and Backbone Cleavage after Aspartate in Intact Proteins by NMR Spectroscopy. *J. Biomol. NMR* **2021**, 1–12.

(670) Aksnes, H.; Ree, R.; Arnesen, T. Co-Translational, Post-Translational, and Non-Catalytic Roles of N-Terminal Acetyltransferases. *Mol. Cell* **2019**, *73*, 1097–1114.

(671) Deng, S.; Marmorstein, R. Protein N-Terminal Acetylation: Structural Basis, Mechanism, Versatility, and Regulation. *Trends Biochem. Sci.* **2021**, *46*, 15–27.

(672) Barbieri, L.; Luchinat, E.; Banci, L. Structural Insights of Proteins in Sub-Cellular Compartments: In-Mitochondria NMR. *Biochim. Biophys. Acta* **2014**, *1843*, 2492–2496.

(673) Mercatelli, E.; Barbieri, L.; Luchinat, E.; Banci, L. Direct Structural Evidence of Protein Redox Regulation Obtained by In-Cell NMR. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2016**, *1863*, 198–204.

(674) Rose, H. M.; Stuiver, M.; Thongwichian, R.; Theillet, F.-X.; Feller, S. M.; Selenko, P. Quantitative NMR Analysis of Erk Activity and Inhibition by U0126 in a Panel of Patient-Derived Colorectal Cancer Cell Lines. *Biochim. Biophys. Acta* **2013**, *1834*, 1396–1401.

(675) Kumar, A.; Gopalswamy, M.; Wishart, C.; Henze, M.; Eschen-Lippold, L.; Donnelly, D.; Balbach, J. N-Terminal Phosphorylation of Parathyroid Hormone (PTH) Abolishes Its Receptor Activity. *ACS Chem. Biol.* **2014**, *9*, 2465–2470.

(676) Kumar, A.; Gopalswamy, M.; Wolf, A.; Brockwell, D. J.; Hatzfeld, M.; Balbach, J. Phosphorylation-Induced Unfolding Regulates P19 ^{INK4d} during the Human Cell Cycle. *Proc. Natl. Acad. Sci.* **2018**, *115*, 3344–3349.

(677) Bourgeois, B.; Gui, T.; Hoogeboom, D.; Hocking, H. G.; Richter, G.; Spreitzer, E.; Viertler, M.; Richter, K.; Madl, T.; Burgering, B. M. T. Multiple Regulatory Intrinsically Disordered Motifs Control FOXO4 Transcription Factor Binding and Function. *Cell Rep.* **2021**, *36*, 109446.

(678) Gebel, J.; Tuppi, M.; Chaikuad, A.; Hötte, K.; Schröder, M.; Schulz, L.; Lohr, F.; Gutfreund, N.; Finke, F.; Henrich, E.; et al. P63 Uses a Switch-like Mechanism to Set the Threshold for Induction of Apoptosis. *Nat. Chem. Biol.* **2020**, *16*, 1078–1086.

(679) Putignano, V.; Rosato, A.; Banci, L.; Andreini, C. MetalPDB in 2018: A Database of Metal Sites in Biological Macromolecular Structures. *Nucleic Acids Res.* **2018**, *46*, D459–D464.

(680) Atrián-Blasco, E.; Gonzalez, P.; Santoro, A.; Alies, B.; Faller, P.; Hureau, C. Cu and Zn Coordination to Amyloid Peptides: From Fascinating Chemistry to Debated Pathological Relevance. *Coord. Chem. Rev.* **2018**, *371*, 38–55.

(681) Nguyen, P. H.; Ramamoorthy, A.; Sahoo, B. R.; Zheng, J.; Faller, P.; Straub, J. E.; Dominguez, L.; Shea, J.-E.; Dokholyan, N. V.; De Simone, A.; et al. Amyloid Oligomers: A Joint Experimental/Computational Perspective on Alzheimer's Disease, Parkinson's Disease, Type II Diabetes, and Amyotrophic Lateral Sclerosis. *Chem. Rev.* **2021**, *121*, 2545–2647.

(682) Waldron, K. J.; Robinson, N. J. How Do Bacterial Cells Ensure That Metalloproteins Get the Correct Metal? *Nat. Rev. Microbiol.* **2009**, *7*, 25–35.

(683) Foster, A. W.; Osman, D.; Robinson, N. J. Metal Preferences and Metallation. *J. Biol. Chem.* **2014**, *289*, 28095–28103.

(684) Yannone, S. M.; Hartung, S.; Menon, A. L.; Adams, M. W.; Tainer, J. A. Metals in Biology: Defining Metalloproteomes. *Curr. Opin. Biotechnol.* **2012**, *23*, 89–95.

(685) New, E. J.; Wimmer, V. C.; Hare, D. J. Promises and Pitfalls of Metal Imaging in Biology. *Cell Chem. Biol.* **2018**, *25*, 7–18.

(686) Wang, H.; Zhou, Y.; Xu, X.; Li, H.; Sun, H. Metalloproteomics in Conjunction with Other Omics for Uncovering the Mechanism of Action of Metallodrugs: Mechanism-Driven New Therapy Development. *Curr. Opin. Chem. Biol.* **2020**, *55*, 171–179.

(687) Jensen, M. R.; Hass, M. A. S.; Hansen, D. F.; Led, J. J. Investigating Metal-Binding in Proteins by Nuclear Magnetic Resonance. *Cell. Mol. Life Sci.* **2007**, *64*, 1085–1104.

(688) Piccioli, M.; Turano, P. Transient Iron Coordination Sites in Proteins: Exploiting the Dual Nature of Paramagnetic NMR. *Coord. Chem. Rev.* **2015**, *284*, 313–328.

(689) Arnesano, F. NMR Spectroscopy to Study the Fate of Metallodrugs in Cells. *Curr. Opin. Chem. Biol.* **2021**, *61*, 214–226.

(690) Giorgi, C.; Danese, A.; Missiroli, S.; Patergnani, S.; Pinton, P. Calcium Dynamics as a Machine for Decoding Signals. *Trends Cell Biol.* **2018**, *28*, 258–273.

(691) Bagur, R.; Hajnóczky, G. Intracellular Ca2+ Sensing: Its Role in Calcium Homeostasis and Signaling. *Mol. Cell* **2017**, *66*, 780–788.

(692) Villalobo, A.; Ishida, H.; Vogel, H. J.; Berchtold, M. W. Calmodulin as a Protein Linker and a Regulator of Adaptor/Scaffold Proteins. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2018**, *1865*, 507–521.

(693) Lian, L.-Y. NMR Studies of Weak Protein-Protein Interactions. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *71*, 59–72.

(694) Yadav, D. K.; Lukavsky, P. J. NMR Solution Structure Determination of Large RNA-Protein Complexes. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *97*, 57–81.

(695) Mishima, M.; Maesaki, R.; Kasa, M.; Watanabe, T.; Fukata, M.; Kaibuchi, K.; Hakoshima, T. Structural Basis for Tubulin Recognition by Cytoplasmic Linker Protein 170 and Its Autoinhibition. *Proc. Natl. Acad. Sci.* **2007**, *104*, 10346–10351.

(696) Shimada, I.; Ueda, T.; Matsumoto, M.; Sakakura, M.; Osawa, M.; Takeuchi, K.; Nishida, N.; Takahashi, H. Cross-Saturation and Transferred Cross-Saturation Experiments. *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *54*, 123–140.

(697) Zhou, X. Z.; Lu, K. P. The Isomerase PIN1 Controls Numerous Cancer-Driving Pathways and Is a Unique Drug Target. *Nat. Rev. Cancer* **2016**, *16*, 463–478.

(698) Wennerstrom, H.; Vallina Estrada, E.; Danielsson, J.; Oliveberg, M. Colloidal Stability of the Living Cell. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 10113–10121.

(699) Purvis, J. E.; Lahav, G. Encoding and Decoding Cellular Information through Signaling Dynamics. *Cell* **2013**, *152*, 945–956.

(700) Liu, Y.; Beyer, A.; Aebersold, R. On the Dependency of Cellular Protein Levels on MRNA Abundance. *Cell* **2016**, *165*, 535–550.

(701) Shis, D. L.; Bennett, M. R.; Igoshin, O. A. Dynamics of Bacterial Gene Regulatory Networks. *Annu. Rev. Biophys.* **2018**, *47*, 447–467.

(702) Blank, H. M.; Papoulas, O.; Maitra, N.; Garge, R.; Kennedy, B. K.; Schilling, B.; Marcotte, E. M.; Polymenis, M. Abundances of Transcripts, Proteins, and Metabolites in the Cell Cycle of Budding Yeast Reveal Coordinate Control of Lipid Metabolism. *Mol. Biol. Cell* **2020**, *31*, 1069–1084.

(703) Mahdessian, D.; Cesnik, A. J.; Gnann, C.; Danielsson, F.; Stenström, L.; Arif, M.; Zhang, C.; Le, T.; Johansson, F.; Shutten, R.; et al. Spatiotemporal Dissection of the Cell Cycle with Single-Cell Proteogenomics. *Nature* **2021**, *590*, 649–654.

(704) Guilhas, B.; Walter, J.-C.; Rech, J.; David, G.; Walliser, N. O.; Palmeri, J.; Mathieu-Demaziere, C.; Parmeggiani, A.; Bouet, J.-Y.; Le Gall, A.; et al. ATP-Driven Separation of Liquid Phase Condensates in Bacteria. *Mol. Cell* **2020**, *79*, 293-303.e4.

(705) Helm, M. S.; Dankovich, T. M.; Mandad, S.; Rammner, B.; Jähne, S.; Salimi, V.; Koerbs, C.; Leibrandt, R.; Urlaub, H.; Schikorski, T.; et al. A Large-Scale Nanoscopy and Biochemistry Analysis of Postsynaptic Dendritic Spines. *Nat. Neurosci.* **2021**, *24*, 1151–1162.

(706) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular Condensates: Organizers of Cellular Biochemistry. *Nat. Rev. Mol. Cell Biol.* **2017**, 1–14.

(707) Xiang, Y.; Surovtsev, I. V.; Chang, Y.; Govers, S. K.; Parry, B. R.; Liu, J.; Jacobs-Wagner, C. Interconnecting Solvent Quality, Transcription, and Chromosome Folding in Escherichia Coli. *Cell* **2021**, *184*, 3626-3642.e14.

(708) Murakami, K.; Kajimoto, S.; Shibata, D.; Kuroi, K.; Fujii, F.; Nakabayashi, T. Observation of Liquid–Liquid Phase Separation of Ataxin-3 and Quantitative Evaluation of Its Concentration in a Single Droplet Using Raman Microscopy. *Chem. Sci.* **2021**, *12*, 7411–7418.

(709) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty Years on: The Impact of Fragments on Drug Discovery. *Nat. Rev. Drug Discov.* **2016**, *15*, 605–619.

(710) Osborne, J.; Panova, S.; Rapti, M.; Urushima, T.; Jhoti, H. Fragments: Where Are We Now? *Biochem. Soc. Trans.* **2020**, *48*, 271–280.

(711) Horst, R.; Farley, K. A.; Kormos, B. L.; Withka, J. M. NMR Spectroscopy: The Swiss Army Knife of Drug Discovery. *J. Biomol. NMR* **2020**, 1–11.

(712) LeBlanc, R. M.; Mesleh, M. F. A Drug Discovery Toolbox for Nuclear Magnetic Resonance (NMR) Characterization of Ligands and Their Targets. *Drug Discov. Today Technol.* **2020**, 1–10.

(713) Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; et al. A Comprehensive Map of Molecular Drug Targets. *Nat. Rev. Drug Discov.* **2017**, *16*, 19–34.

(714) Vasile, F.; Panigada, M.; Siccardi, A.; Potenza, D.; Tiana, G. A Combined NMR-Computational Study of the Interaction between Influenza Virus Hemagglutinin and Sialic Derivatives from Human and Avian Receptors on the Surface of Transfected Cells. *Int. J. Mol. Sci.* **2018**, *19*.

(715) Potenza, D.; Belvisi, L. Transferred-NOE NMR Experiments on Intact Human Platelets: Receptor-Bound Conformation of RGD-Peptide Mimics. *Org. Biomol. Chem.* **2008**, *6*, 258–262.

(716) Nardelli, F.; Paissoni, C.; Quilici, G.; Gori, A.; Traversari, C.; Valentinis, B.; Sacchi, A.; Corti, A.; Curnis, F.; Ghitti, M.; et al. Succinimide-Based Conjugates Improve IsoDGR Cyclopeptide Affinity to Avβ3 without Promoting Integrin Allosteric Activation. *J. Med. Chem.* **2018**, *61*, 7474–7485.

(717) Farina, B.; de Paola, I.; Russo, L.; Capasso, D.; Liguoro, A.; Gatto, A. D.; Saviano, M.; Pedone, P. V.; Di Gaetano, S.; Malgieri, G.; et al. A Combined NMR and Computational Approach to Determine the RGDechi-HCit-Av B3 Integrin Recognition Mode in Isolated Cell Membranes. *Chem. - Eur. J.* **2016**, *22*, 681–693.

(718) Russo, L.; Farina, B.; Del Gatto, A.; Comegna, D.; Di Gaetano, S.; Capasso, D.; Liguoro, A.; Malgieri, G.; Saviano, M.; Fattorusso, R.; et al. Deciphering RGDechi Peptide-α 5β 1integrin Interaction Mode in Isolated Cell Membranes. *Pept. Sci.* **2018**, *110*, e24065-14.

(719) Assadi-Porter, F. M.; Tonelli, M.; Maillet, E.; Hallenga, K.; Benard, O.; Max, M.; Markley, J. L. Direct NMR Detection of the Binding of Functional Ligands to the Human Sweet Receptor, a Heterodimeric Family 3 GPCR. *J. Am. Chem. Soc.* **2008**, *130*, 7212–7213.

(720) Assadi-Porter, F. M.; Tonelli, M.; Maillet, E. L.; Markley, J. L.; Max, M. Interactions between the Human Sweet-Sensing T1R2-T1R3 Receptor and Sweeteners Detected by Saturation Transfer Difference NMR Spectroscopy. *Biochim. Biophys. Acta* **2010**, *1798*, 82–86.

(721) Bartoschek, S.; Klabunde, T.; Defossa, E.; Dietrich, V.; Stengelin, S.; Griesinger, C.; Carlomagno, T.; Focken, I.; Wendt, K. U. Drug Design for G-Protein-Coupled Receptors by a Ligand-Based NMR Method. *Angew. Chem. Int. Ed Engl.* **2010**, *49*, 1426–1429.

(722) Cox, B. D.; Mehta, A. K.; DiRaddo, J. O.; Liotta, D. C.; Wilson, L. J.; Snyder, J. P. Structural Analysis of CXCR4 - Antagonist Interactions Using Saturation-Transfer Double-Difference NMR. *Biochem. Biophys. Res. Commun.* **2015**, *466*, 28–32.

(723) Nikolaev, Y. V.; Kochanowski, K.; Link, H.; Sauer, U.; Allain, F. H.-T. Systematic Identification of Protein-Metabolite Interactions in Complex Metabolite Mixtures by Ligand-Detected Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2016**, *55*, 2590–2600.

(724) Diether, M.; Nikolaev, Y.; Allain, F. H.; Sauer, U. Systematic Mapping of Protein-Metabolite Interactions in Central Metabolism of Escherichia Coli. *Mol. Syst. Biol.* **2019**, *15*, e9008.

(725) Strasser, A.; Wittmann, H.-J.; Seifert, R. Binding Kinetics and Pathways of Ligands to GPCRs. *Trends Pharmacol. Sci.* **2017**, *38*, 717–732.

(726) Georgi, V.; Schiele, F.; Berger, B.-T.; Steffen, A.; Marin Zapata, P. A.; Briem, H.; Menz, S.; Preusse, C.; Vasta, J. D.; Robers, M. B.; et al. Binding Kinetics Survey of the Drugged Kinome. *J. Am. Chem. Soc.* **2018**, *140*, 15774–15782.

(727) Stefaniak, J.; Huber, K. V. M. Importance of Quantifying Drug-Target Engagement in Cells. *ACS Med. Chem. Lett.* **2020**, *11*, 403–406.

(728) Wishart, D. S. Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol. Rev.* **2019**, *99*, 1819–1875.

(729) Vignoli, A.; Ghini, V.; Meoni, G.; Licari, C.; Takis, P. G.; Tenori, L.; Turano, P.; Luchinat, C. High-Throughput Metabolomics by 1D NMR. *Angew. Chem. Int. Ed Engl.* **2019**, *58*, 968–994.

(730) Xu, Z.; Liu, C.; Zhao, S.; Chen, S.; Zhao, Y. Molecular Sensors for NMR-Based Detection. *Chem. Rev.* **2019**, *119*, 195–230.

(731) Tressler, C. M.; Zondlo, N. J. Perfluoro-Tert-Butyl Hydroxyprolines as Sensitive, Conformationally Responsive Molecular Probes: Detection of Protein Kinase Activity by 19F NMR. *ACS Chem. Biol.* **2020**, *15*, 1096–1103.

(732) Dugger, S. A.; Platt, A.; Goldstein, D. B. Drug Development in the Era of Precision Medicine. *Nat. Rev. Drug Discov.* **2018**, *17*, 183–196.

(733) Nussinov, R.; Jang, H.; Nir, G.; Tsai, C.-J.; Cheng, F. A New Precision Medicine Initiative at the Dawn of Exascale Computing. *Signal Transduct. Target. Ther.* **2021**, *6*, 3.

(734) Murugan, A. K.; Grieco, M.; Tsuchida, N. RAS Mutations in Human Cancers: Roles in Precision Medicine. *Semin. Cancer Biol.* **2019**, *59*, 23–35.

(735) Bonucci, A.; Ouari, O.; Guigliarelli, B.; Belle, V.; Mileo, E. In-Cell EPR: Progress towards Structural Studies Inside Cells. *ChemBioChem* **2019**, *21*, 451–460.

(736) Bieber, A.; Drescher, M. Chapter 10:In-Cell EPR. In *In-cell NMR Spectroscopy*; 2019; pp 152–169.

(737) Torricella, F.; Pierro, A.; Mileo, E.; Belle, V.; Bonucci, A. Nitroxide Spin Labels and EPR Spectroscopy: A Powerful Association for Protein Dynamics Studies. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2021**, *1869*, 140653.

(738) Schmidt, M. J.; Borbas, J.; Drescher, M.; Summerer, D. A Genetically Encoded Spin Label for Electron Paramagnetic Resonance Distance Measurements. *J. Am. Chem. Soc.* **2014**, *136*, 1238–1241.

(739) Kucher, S.; Korneev, S.; Klare, J. P.; Klose, D.; Steinhoff, H.-J. *In Cell* Gd $^{3+}$ -Based Site-Directed Spin Labeling and EPR Spectroscopy of EGFP. *Phys. Chem. Chem. Phys.* **2020**, *22*, 13358–13362.

(740) Braun, T. S.; Stehle, J.; Kacprzak, S.; Carl, P.; Höfer, P.; Subramaniam, V.; Drescher, M. Intracellular Protein–Lipid Interactions Studied by Rapid-Scan Electron Paramagnetic Resonance Spectroscopy. *J. Phys. Chem. Lett.* **2021**, *12*, 2471–2475.

(741) Chow, G. K.; Chavan, A. G.; Heisler, J. C.; Chang, Y.-G.; LiWang, A.; Britt, R. D. Monitoring Protein–Protein Interactions in the Cyanobacterial Circadian Clock in Real Time via Electron Paramagnetic Resonance Spectroscopy. *Biochemistry* **2020**, *59*, 2387–2400.

(742) Joseph, B.; Jaumann, E. A.; Sikora, A.; Barth, K.; Prisner, T. F.; Cafiso, D. S. In Situ Observation of Conformational Dynamics and Protein Ligand–Substrate Interactions in Outer-Membrane Proteins with DEER/PELDOR Spectroscopy. *Nat. Protoc.* **2019**, *14*, 2344–2369.

(743) Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446.

(744) Goldfarb, D. Pulse EPR in Biological Systems – Beyond the Expert's Courtyard. *J. Magn. Reson.* **2019**, *306*, 102–108.

(745) Schiemann, O.; Heubach, C. A.; Abdullin, D.; Ackermann, K.; Azarkh, M.; Bagryanskaya, E. G.; Drescher, M.; Endeward, B.; Freed, J. H.; Galazzo, L.; et al. Benchmark Test and Guidelines for DEER/PELDOR Experiments on Nitroxide-Labeled Biomolecules. *J. Am. Chem. Soc.* **2021**, *143*, 17875–17890.

(746) Hasanbasri, Z.; Singewald, K.; Gluth, T. D.; Driesschaert, B.; Saxena, S. Cleavage-Resistant Protein Labeling With Hydrophilic Trityl Enables Distance Measurements *In-Cell*. *J. Phys. Chem. B* **2021**, *125*, 5265–5274.

(747) Collauto, A.; Bülow, S.; Gophane, D. B.; Saha, S.; Stelzl, L. S.; Hummer, G.; Sigurdsson, S. T.; Prisner, T. F. Compaction of RNA Duplexes in the Cell**. *Angew. Chem. Int. Ed.* **2020**, *59*, 23025–23029.

(748) Fleck, N.; Heubach, C. A.; Hett, T.; Haege, F. R.; Bawol, P. P.; Baltruschat, H.; Schiemann, O. SLIM: A Short-Linked, Highly Redox-Stable Trityl Label for High-Sensitivity In-Cell EPR Distance Measurements. *Angew. Chem. Int. Ed.* **2020**, *59*, 9767–9772.

(749) Miao, Q.; Zurlo, E.; Bruin, D.; Wondergem, J. A. J.; Timmer, M.; Blok, A.; Heinrich, D.; Overhand, M.; Huber, M.; Ubbink, M. A Two-Armed Probe for In-Cell DEER Measurements on Proteins**. *Chem. – Eur. J.* **2020**, *26*, 17128–17133.

(750) Worswick, S. G.; Spencer, J. A.; Jeschke, G.; Kuprov, I. Deep Neural Network Processing of DEER Data. *Sci. Adv.* **2018**, *4*, eaat5218.

(751) Sweger, S. R.; Pribitzer, S.; Stoll, S. Bayesian Probabilistic Analysis of DEER Spectroscopy Data Using Parametric Distance Distribution Models. *J. Phys. Chem. A* **2020**, *124*, 6193–6202.

(752) del Alamo, D.; Jagessar, K. L.; Meiler, J.; Mchaourab, H. S. Methodology for Rigorous Modeling of Protein Conformational Changes by Rosetta Using DEER Distance Restraints. *PLOS Comput. Biol.* **2021**, *17*, e1009107.

(753) Jeschke, G. Conformational Dynamics and Distribution of Nitroxide Spin Labels. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *72*, 42–60.

(754) Klose, D.; Holla, A.; Gmeiner, C.; Nettels, D.; Ritsch, I.; Bross, N.; Yulikov, M.; Allain, F. H.-T.; Schuler, B.; Jeschke, G. Resolving Distance Variations by Single-Molecule FRET and EPR Spectroscopy Using Rotamer Libraries. *Biophys. J.* **2021**, S0006349521007554.

(755) Azarkh, M.; Okle, O.; Eyring, P.; Dietrich, D. R.; Drescher, M. Evaluation of Spin Labels for In-Cell EPR by Analysis of Nitroxide Reduction in Cell Extract of Xenopus Laevis Oocytes. *J. Magn. Reson.* **2011**, *212*, 450–454.

(756) Paletta, J. T.; Pink, M.; Foley, B.; Rajca, S.; Rajca, A. Synthesis and Reduction Kinetics of Sterically Shielded Pyrrolidine Nitroxides. *Org. Lett.* **2012**, *14*, 5322–5325.

(757) Lawless, M. J.; Shimshi, A.; Cunningham, T. F.; Kinde, M. N.; Tang, P.; Saxena, S. Analysis of Nitroxide-Based Distance Measurements in Cell Extracts and in Cells by Pulsed ESR Spectroscopy. *ChemPhysChem* **2017**, *18*, 1653–1660.

(758) Ketter, S.; Gopinath, A.; Rogozhnikova, O.; Trukhin, D.; Tormyshev, V. M.; Bagryanskaya, E. G.; Joseph, B. In Situ Labeling and Distance Measurements of Membrane Proteins in *E. Coli* Using Finland and OX063 Trityl Labels. *Chem. – Eur. J.* **2021**, *27*, 2299–2304.

(759) Jassoy, J. J.; Berndhäuser, A.; Duthie, F.; Kühn, S. P.; Hagelueken, G.; Schiemann, O. Versatile Trityl Spin Labels for Nanometer Distance Measurements on Biomolecules In Vitro and within Cells. *Angew. Chem. Int. Ed.* **2017**, *56*, 177–181.

(760) Feintuch, A.; Otting, G.; Goldfarb, D. Gd3+ Spin Labeling for Measuring Distances in Biomacromolecules. In *Methods in Enzymology*; Elsevier, 2015; Vol. 563, pp 415–457.

(761) Martorana, A.; Bellapadrona, G.; Feintuch, A.; Di Gregorio, E.; Aime, S.; Goldfarb, D. Probing Protein Conformation in Cells by EPR Distance Measurements Using Gd 3+Spin Labeling. *J. Am. Chem. Soc.* **2014**, *136*, 13458–13465.

(762) Azarkh, M.; Bieber, A.; Qi, M.; Fischer, J. W. A.; Yulikov, M.; Godt, A.; Drescher, M. Gd(III)–Gd(III) Relaxation-Induced Dipolar Modulation Enhancement for In-Cell Electron Paramagnetic Resonance Distance Determination. *J. Phys. Chem. Lett.* **2019**, *10*, 1477–1481.

(763) Joseph, B.; Sikora, A.; Bordignon, E.; Jeschke, G.; Cafiso, D. S.; Prisner, T. F. Distance Measurement on an Endogenous Membrane Transporter in E. Coli Cells and Native Membranes Using EPR Spectroscopy. *Angew. Chem. Int. Ed.* **2015**, *54*, 6196–6199.

(764) Nilaweera, T. D.; Nyenhuis, D. A.; Nakamoto, R. K.; Cafiso, D. S. Disulfide Chaperone Knockouts Enable In Vivo Double Spin Labeling of an Outer Membrane Transporter. *Biophys. J.* **2019**, *117*, 1476–1484.

(765) Nilaweera, T. D.; Nyenhuis, D. A.; Cafiso, D. S. Structural Intermediates Observed Only in Intact Escherichia Coli Indicate a Mechanism for TonB-Dependent Transport. *eLife* **2021**, *10*, e68548.

(766) Mascali, F. C.; Ching, H. Y. V.; Rasia, R. M.; Un, S.; Tabares, L. C. Using Genetically Encodable Self-Assembling Gd ^{III} Spin Labels To Make In-Cell Nanometric Distance Measurements. *Angew. Chem. Int. Ed.* **2016**, *55*, 11041–11043.

(767) Schmidt, M. J.; Fedoseev, A.; Summerer, D.; Drescher, M. Genetically Encoded Spin Labels for In Vitro and In-Cell EPR Studies of Native Proteins. In *Methods in Enzymology*; Elsevier, 2015; Vol. 563, pp 483–502.

(768) Widder, P.; Schuck, J.; Summerer, D.; Drescher, M. Combining Site-Directed Spin Labeling *in Vivo* and in-Cell EPR Distance Determination. *Phys. Chem. Chem. Phys.* **2020**, *22*, 4875–4879.

(769) Galazzo, L.; Meier, G.; Timachi, M. H.; Hutter, C. A. J.; Seeger, M. A.; Bordignon, E. Spin-Labeled Nanobodies as Protein Conformational Reporters for Electron Paramagnetic Resonance in Cellular Membranes. *Proc. Natl. Acad. Sci.* **2020**, *117*, 2441–2448.

(770) Igarashi, R.; Sakai, T.; Hara, H.; Tenno, T.; Tanaka, T.; Tochio, H.; Shirakawa, M. Distance Determination in Proteins inside *Xenopus Laevis* Oocytes by Double Electron−Electron Resonance Experiments. *J. Am. Chem. Soc.* **2010**, *132*, 8228–8229.

(771) Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H. Effect of Freezing Conditions on Distances and Their Distributions Derived from

Double Electron Electron Resonance (DEER): A Study of Doubly-Spin-Labeled T4 Lysozyme. *J. Magn. Reson.* **2012**, *216*, 69–77.

(772) Schmidt, T.; Jeon, J.; Okuno, Y.; Chiliveri, S. C.; Clore, G. M. Submillisecond Freezing Permits Cryoprotectant-Free EPR Double Electron−Electron Resonance Spectroscopy. *ChemPhysChem* **2020**, *21*, 1224–1229.

(773) Lerner, E.; Cordes, T.; Ingargiola, A.; Alhadid, Y.; Chung, S.; Michalet, X.; Weiss, S. Toward Dynamic Structural Biology: Two Decades of Single-Molecule Förster Resonance Energy Transfer. *Science* **2018**, *359*, eaan1133.

(774) Ebbinghaus, S.; Dhar, A.; McDonald, J. D.; Gruebele, M. Protein Folding Stability and Dynamics Imaged in a Living Cell. *Nat. Methods* **2010**, *7*, 319–323.

(775) Wirth, A. J.; Platkov, M.; Gruebele, M. Temporal Variation of a Protein Folding Energy Landscape in the Cell. *J. Am. Chem. Soc.* **2013**, *135*, 19215–19221.

(776) Sukenik, S.; Ren, P.; Gruebele, M. Weak Protein–Protein Interactions in Live Cells Are Quantified by Cell-Volume Modulation. *Proc. Natl. Acad. Sci.* **2017**, 201700818.

(777) Guin, D.; Gelman, H.; Wang, Y.; Gruebele, M. Heat Shock-Induced Chaperoning by Hsp70 Is Enabled in-Cell. *PLOS ONE* **2019**, *14*, e0222990.

(778) Feng, R.; Gruebele, M.; Davis, C. M. Quantifying Protein Dynamics and Stability in a Living Organism. *Nat. Commun.* **2019**, *10*, 1179.

(779) Gnutt, D.; Timr, S.; Ahlers, J.; König, B.; Manderfeld, E.; Heyden, M.; Sterpone, F.; Ebbinghaus, S. Stability Effect of Quinary Interactions Reversed by Single Point Mutations. *J Am Chem Soc* **2019**, 10.

(780) Guin, D.; Gruebele, M. Chaperones Hsc70 and Hsp70 Bind to the Protein PGK Differently inside Living Cells. *J. Phys. Chem. B* **2020**, *124*, 3629–3635.

(781) Mouton, S. N.; Thaller, D. J.; Crane, M. M.; Rempel, I. L.; Terpstra, O. T.; Steen, A.; Kaeberlein, M.; Lusk, C. P.; Boersma, A. J.; Veenhoff, L. M. A Physicochemical Perspective of Aging from Single-Cell Analysis of PH, Macromolecular and Organellar Crowding in Yeast. *eLife* **2020**, *9*, 959–23.

(782) Basak, S.; Sakia, N.; Dougherty, L.; Guo, Z.; Wu, F.; Mindlin, F.; Lary, J. W.; Cole, J. L.; Ding, F.; Bowen, M. E. Probing Interdomain Linkers and Protein Supertertiary Structure In Vitro and in Live Cells with Fluorescent Protein Resonance Energy Transfer. *J. Mol. Biol.* **2021**, *433*, 166793.

(783) Sakon, J. J.; Weninger, K. R. Detecting the Conformation of Individual Proteins in Live Cells. *Nat. Methods* **2010**, *7*, 203–205.

(784) König, I.; Zarrine-Afsar, A.; Aznauryan, M.; Soranno, A.; Wunderlich, B.; Dingfelder, F.; Stüber, J. C.; Plückthun, A.; Nettels, D.; Schuler, B. Single-Molecule Spectroscopy of Protein Conformational Dynamics in Live Eukaryotic Cells. *Nat. Methods* **2015**, *12*, 773–779.

(785) Okamoto, K.; Hibino, K.; Sako, Y. In-Cell Single-Molecule FRET Measurements Reveal Three Conformational State Changes in RAF Protein. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2020**, *1864*, 129358.

(786) König, I.; Soranno, A.; Nettels, D.; Schuler, B. Impact of In-Cell and In-Vitro Crowding on the Conformations and Dynamics of an Intrinsically Disordered Protein. *Angew. Chem. Int. Ed.* **2021**, *60*, 10724–10729.

(787) Asher, W. B.; Geggier, P.; Holsey, M. D.; Gilmore, G. T.; Pati, A. K.; Meszaros, J.; Terry, D. S.; Mathiasen, S.; Kaliszewski, M. J.; McCauley, M. D.; et al. Single-Molecule FRET Imaging of GPCR Dimers in Living Cells. *Nat. Methods* **2021**, *18*, 397–405.

(788) Phillip, Y.; Kiss, V.; Schreiber, G. Protein-Binding Dynamics Imaged in a Living Cell. *Proc. Natl. Acad. Sci.* **2012**, *109*, 1461–1466.

(789) Davis, C. M.; Gruebele, M. Cellular Sticking Can Strongly Reduce Complex Binding by Speeding Dissociation. *J. Phys. Chem. B* **2021**, *125*, 3815–3823.

(790) Lerner, E.; Barth, A.; Hendrix, J.; Ambrose, B.; Birkedal, V.; Blanchard, S. C.; Börner, R.; Sung Chung, H.; Cordes, T.; Craggs, T. D.; et al. FRET-Based Dynamic Structural Biology: Challenges, Perspectives and an Appeal for Open-Science Practices. *eLife* **2021**, *10*, e60416.

(791) Schuler, B. Perspective: Chain Dynamics of Unfolded and Intrinsically Disordered Proteins from Nanosecond Fluorescence Correlation Spectroscopy Combined with Single-Molecule FRET. *J. Chem. Phys.* **2018**, *149*, 010901.

(792) Dimura, M.; Peulen, T. O.; Hanke, C. A.; Prakash, A.; Gohlke, H.; Seidel, C. A. Quantitative FRET Studies and Integrative Modeling Unravel the Structure and Dynamics of Biomolecular Systems. *Curr. Opin. Struct. Biol.* **2016**, *40*, 163–185.

(793) Dimura, M.; Peulen, T.-O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H. Automated and Optimally FRET-Assisted Structural Modeling. *Nat. Commun.* **2020**, *11*, 5394.

(794) Quast, R. B.; Margeat, E. Single-Molecule FRET on Its Way to Structural Biology in Live Cells. *Nat. Methods* **2021**, *18*, 344–345.

(795) Sánchez-Rico, C.; Voith von Voithenberg, L.; Warner, L.; Lamb, D. C.; Sattler, M. Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by SpFRET and NMR Spectroscopy. *Chem. - Eur. J.* **2017**, *23*, 14267–14277.

(796) Liu, B.; Stone, O. J.; Pablo, M.; Herron, J. C.; Nogueira, A. T.; Dagliyan, O.; Grimm, J. B.; Lavis, L. D.; Elston, T. C.; Hahn, K. M. Biosensors Based on Peptide Exposure Show Single Molecule Conformations in Live Cells. *Cell* **2021**, *184*, 5670-5685.e23.

(797) Savitski, M. M.; Reinhard, F. B. M.; Franken, H.; Werner, T.; Savitski, M. F.; Eberhard, D.; Molina, D. M.; Jafari, R.; Dovega, R. B.; Klaeger, S.; et al. Tracking Cancer Drugs in Living Cells by Thermal Profiling of the Proteome. *Science* **2014**, *346*, 1255784–12.

(798) Franken, H.; Mathieson, T.; Childs, D.; Sweetman, G. M. A.; Werner, T.; gel, I. T. ouml; Doce, C.; Gade, S.; Bantscheff, M.; Drewes, G.; et al. Thermal Proteome Profiling for Unbiased Identification of Direct and Indirect Drug Targets Using Multiplexed Quantitative Mass Spectrometry. *Nat. Protoc.* **2015**, *10*, 1567–1593.

(799) Ochoa, D.; Jarnuczak, A. F.; itez, C. V. x000E9; Gehre, M.; Soucheray, M.; Mateus, A. x000E9; Kleefeldt, A. A.; Hill, A.; Garcia-Alonso, L.; Stein, F.; et al. The Functional Landscape of the Human Phosphoproteome. *Nat. Biotechnol.* **2019**, 1–14.

(800) Mateus, A.; Hevler, J.; Bobonis, J.; Kurzawa, N.; Shah, M.; Mitosch, K.; Goemans, C. V.; Helm, D.; Stein, F.; Typas, A.; et al. The Functional Proteome Landscape of Escherichia Coli. *Nature* **2020**, *588*, 473–478.

(801) Perrin, J.; Werner, T.; Kurzawa, N.; Rutkowska, A.; Childs, D. D.; Kalxdorf, M.; Poeckel, D.; Stonehouse, E.; Strohmer, K.; Heller, B.; et al. Identifying Drug Targets in Tissues and Whole Blood with Thermal-Shift Profiling. *Nat. Biotechnol.* **2020**, 1–15.

(802) Mateus, A.; Savitski, M. M.; Piazza, I. The Rise of Proteome-wide Biophysics. *Mol. Syst. Biol.* **2021**, *17*.

(803) Feng, Y.; De Franceschi, G.; Kahraman, A.; Soste, M.; Melnik, A.; Boersema, P. J.; de Laureto, P. P.; Nikolaev, Y.; Oliveira, A. P.; Picotti, P. Global Analysis of Protein Structural Changes in Complex Proteomes. *Nat. Biotechnol.* **2014**, 1–13.

(804) Jafari, R.; Almqvist, H.; Axelsson, H.; Ignatushchenko, M.; ck, T. L. auml; Nordlund, P. auml r; Molina, D. M. The Cellular Thermal Shift Assay for Evaluating Drug Target Interactions in Cells. *Nat. Protoc.* **2014**, *9*, 2100–2122.

(805) Huang, J. X.; Lee, G.; Cavanaugh, K. E.; Chang, J. W.; Gardel, M. L.; Moellering, R. E. High Throughput Discovery of Functional Protein Modifications by Hotspot Thermal Profiling. *Nat. Methods* **2019**, 1–13.

(806) Huang, J. X.; Wu, D.; Stein, B. D.; Moellering, R. E. Huang et al. Reply. *Nat. Methods* **2021**, *18*, 763–767.

(807) Reinhard, F. B. M.; Eberhard, D.; Werner, T.; Franken, H.; Childs, D.; Doce, C.; Savitski, M. F.; Huber, W.; Bantscheff, M.; Savitski, M. M.; et al. Thermal Proteome Profiling Monitors Ligand Interactions with Cellular Membrane Proteins. *Nat. Methods* **2015**, *12*, 1129–1131.

(808) Martinez Molina, D.; Nordlund, P. The Cellular Thermal Shift Assay: A Novel Biophysical Assay for In Situ Drug Target

Engagement and Mechanistic Biomarker Studies. *Annu. Rev. Pharmacol. Toxicol.* **2016**, *56*, 141–161.

(809) Kurzawa, N.; Becher, I.; Sridharan, S.; Franken, H.; Mateus, A.; Anders, S.; Bantscheff, M.; Huber, W.; Savitski, M. M. A Computational Method for Detection of Ligand-Binding Proteins from Dose Range Thermal Proteome Profiles. *Nat. Commun.* **2020**, *11*, 5783.

(810) Kalxdorf, M.; Günthner, I.; Becher, I.; Kurzawa, N.; Knecht, S.; Savitski, M. M.; Eberl, H. C.; Bantscheff, M. Cell Surface Thermal Proteome Profiling Tracks Perturbations and Drug Targets on the Plasma Membrane. *Nat. Methods* **2021**, *18*, 84–91.

(811) Dai, L.; Prabhu, N.; Yu, L. Y.; Bacanu, S.; Ramos, A. D.; Nordlund, P. Horizontal Cell Biology: Monitoring Global Changes of Protein Interaction States with the Proteome-Wide Cellular Thermal Shift Assay (CETSA). *Annu. Rev. Biochem.* **2019**, *88*, 383–408.

(812) Becher, I.; Andrés-Pons, A.; Romanov, N.; Stein, F.; Schramm, M.; Baudin, F.; Helm, D.; Kurzawa, N.; Mateus, A.; Mackmull, M.-T.; et al. Pervasive Protein Thermal Stability Variation during the Cell Cycle. *Cell* **2018**, *173*, 1495-1507.e18.

(813) Tsukidate, T.; Li, Q.; Hang, H. C. Targeted and Proteome-Wide Analysis of Metabolite-Protein Interactions. *Curr. Opin. Chem. Biol.* **2020**, *54*, 19–27.

(814) Mateus, A.; Kurzawa, N.; Becher, I.; Sridharan, S.; Helm, D.; Stein, F.; Typas, A.; Savitski, M. M. Thermal Proteome Profiling for Interrogating Protein Interactions. *Mol. Syst. Biol.* **2020**, *16*, 2019.2012.2030.-11.

(815) Prabhu, N.; Dai, L.; Nordlund, P. CETSA in Integrated Proteomics Studies of Cellular Processes. *Curr. Opin. Chem. Biol.* **2020**, *54*, 54–62.

(816) Liu, X. R.; Zhang, M. M.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chem. Rev.* **2020**, *120*, 4355–4454.

(817) Chavez, J. D.; Wippel, H. H.; Tang, X.; Keller, A.; Bruce, J. E. In-Cell Labeling and Mass Spectrometry for Systems-Level Structural Biology. *Chem. Rev.* **2021**, acs.chemrev.1c00223.

(818) Narayanan, S.; Mitra, G.; Muralidharan, M.; Mathew, B.; Mandal, A. K. Protein Structure–Function Correlation in Living Human Red Blood Cells Probed by Isotope Exchange-Based Mass Spectrometry. *Anal. Chem.* **2015**, *87*, 11812–11818.

(819) Shen, G.; Li, S.; Cui, W.; Liu, S.; Yang, Y.; Gross, M.; Li, W. Membrane Protein Structure in Live Cells: Methodology for Studying Drug Interaction by Mass Spectrometry-Based Footprinting. *Biochemistry* **2018**, *57*, 286–294.

(820) Bamberger, C.; Pankow, S.; Martínez-Bartolomé, S.; Ma, M.; Diedrich, J.; Rissman, R. A.; Yates, J. R. Protein Footprinting via Covalent Protein Painting Reveals Structural Changes of the Proteome in Alzheimer's Disease. *J. Proteome Res.* **2021**, *20*, 2762–2771.

(821) McKenzie-Coe, A.; Montes, N. S.; Jones, L. M. Hydroxyl Radical Protein Footprinting: A Mass Spectrometry-Based Structural Method for Studying the Higher Order Structure of Proteins. *Chem. Rev.* **2021**, acs.chemrev.1c00432.

(822) Li, K. S.; Shi, L.; Gross, M. L. Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization. *Acc. Chem. Res.* **2018**, *51*, 736–744.

(823) Espino, J. A.; Mali, V. S.; Jones, L. M. In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Anal. Chem.* **2015**, *87*, 7971–7978.

(824) Johnson, D. T.; Punshon-Smith, B.; Espino, J. A.; Gershenson, A.; Jones, L. M. Implementing In-Cell Fast Photochemical Oxidation of Proteins in a Platform Incubator with a Movable XY Stage. *Anal. Chem.* **2020**, *92*, 1691–1696.

(825) McKenzie-Coe, A. A.; Johnson, D. T.; Peacock, R. B.; Zhang, Z.; Jones, L. M. Evaluating the Sulfate Radical Anion as a New Reagent for In-Cell Fast Photochemical Oxidation of Proteins. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 1644–1647.

(826) Kaur, U.; Johnson, D. T.; Jones, L. M. Validation of the Applicability of In-Cell Fast Photochemical Oxidation of Proteins across Multiple Eukaryotic Cell Lines. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1372–1379.

(827) Espino, J. A.; Jones, L. M. Illuminating Biological Interactions with in Vivo Protein Footprinting. *Anal. Chem.* **2019**, *91*, 6577–6584.

(828) Espino, J. A.; Zhang, Z.; Jones, L. M. Chemical Penetration Enhancers Increase Hydrogen Peroxide Uptake in *C. Elegans* for *In Vivo* Fast Photochemical Oxidation of Proteins. *J. Proteome Res.* **2020**, *19*, 3708–3715.

(829) Biehn, S. E.; Lindert, S. Accurate Protein Structure Prediction with Hydroxyl Radical Protein Footprinting Data. *Nat. Commun.* **2021**, *12*, 341.

(830) Zhu, Y.; Serra, A.; Guo, T.; Park, J. E.; Zhong, Q.; Sze, S. K. Application of Nanosecond Laser Photolysis Protein Footprinting to Study EGFR Activation by EGF in Cells. *J. Proteome Res.* **2017**, *16*, 2282–2293.

(831) Liu, X. R.; Rempel, D. L.; Gross, M. L. Protein Higher-Order-Structure Determination by Fast Photochemical Oxidation of Proteins and Mass Spectrometry Analysis. *Nat. Protoc.* **2020**, *15*, 3942–3970.

(832) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; et al. Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **2017**, *168*, 527-541.e29.

(833) Wang, Y.; Dix, M. M.; Bianco, G.; Remsberg, J. R.; Lee, H.-Y.; Kalocsay, M.; Gygi, S. P.; Forli, S.; Vite, G.; Lawrence, R. M.; et al. Expedited Mapping of the Ligandable Proteome Using Fully Functionalized Enantiomeric Probe Pairs. *Nat. Chem.* **2019**, 1–14.

(834) Burton, N. R.; Kim, P.; Backus, K. M. Photoaffinity Labelling Strategies for Mapping the Small Molecule–Protein Interactome. *Org. Biomol. Chem.* **2021**, *19*, 7792–7809.

(835) Abbasov, M. E.; Kavanagh, M. E.; Ichu, T.-A.; Lazear, M. R.; Tao, Y.; Crowley, V. M.; am Ende, C. W.; Hacker, S. M.; Ho, J.; Dix, M. M.; et al. A Proteome-Wide Atlas of Lysine-Reactive Chemistry. *Nat. Chem.* **2021**, *13*, 1081–1092.

(836) Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; et al. Reimagining High-Throughput Profiling of Reactive Cysteines for Cell-Based Screening of Large Electrophile Libraries. *Nat. Biotechnol.* **2021**, *39*, 630–641.

(837) Rogawski, R.; Rogel, A.; Bloch, I.; Gal, M.; Horovitz, A.; London, N.; Sharon, M. Intracellular Protein–Drug Interactions Probed by Direct Mass Spectrometry of Cell Lysates. *Angew. Chem. Int. Ed.* **2021**, *60*, 19637–19642.

(838) Rogawski, R.; Sharon, M. Characterizing Endogenous Protein Complexes with Biological Mass Spectrometry. *Chem. Rev.* **2021**, acs.chemrev.1c00217.

(839) Qin, W.; Cho, K. F.; Cavanagh, P. E.; Ting, A. Y. Deciphering Molecular Interactions by Proximity Labeling. *Nat. Methods* **2021**, 1–11.

(840) Huang, J. X.; Coukos, J. S.; Moellering, R. E. Interaction Profiling Methods to Map Protein and Pathway Targets of Bioactive Ligands. *Curr. Opin. Chem. Biol.* **2020**, *54*, 76–84.

(841) Minde, D.-P.; Ramakrishna, M.; Lilley, K. S. Biotin Proximity Tagging Favours Unfolded Proteins and Enables the Study of Intrinsically Disordered Regions. *Commun. Biol.* **2020**, 1–13.

(842) O'Reilly, F. J.; Xue, L.; Graziadei, A.; Sinn, L.; Lenz, S.; Tegunov, D.; Blötz, C.; Singh, N.; Hagen, W. J. H.; Cramer, P.; et al. In-Cell Architecture of an Actively Transcribing-Translating Expressome. *Science* **2020**, *369*, 554–557.

(843) Chavez, J. D.; Schweppe, D. K.; Eng, J. K.; Bruce, J. E. In Vivo Conformational Dynamics of Hsp90 and Its Interactors. *Cell Chem. Biol.* **2016**, *23*, 716–726.

(844) Chavez, J. D.; Mohr, J. P.; Mathay, M.; Zhong, X.; Keller, A.; Bruce, J. E. Systems Structural Biology Measurements by in Vivo Cross-Linking with Mass Spectrometry. *Nat. Protoc.* **2019**, *14*, 2318–2343.

(845) Rey, M.; Dhenin, J.; Kong, Y.; Nouchikian, L.; Filella, I.; Duchateau, M.; Dupré, M.; Pellarin, R.; Duménil, G.; Chamot-

Rooke, J. Advanced *In Vivo* Cross-Linking Mass Spectrometry Platform to Characterize Proteome-Wide Protein Interactions. *Anal. Chem.* **2021**, *93*, 4166–4174.

(846) Wheat, A.; Yu, C.; Wang, X.; Burke, A. M.; Chemmama, I. E.; Kaake, R. M.; Baker, P.; Rychnovsky, S. D.; Yang, J.; Huang, L. Protein Interaction Landscapes Revealed by Advanced in Vivo Cross-Linking–Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2021**, *118*, e2023360118.

(847) Ryl, P. S. J.; Bohlke-Schneider, M.; Lenz, S.; Fischer, L.; Budzinski, L.; Stuiver, M.; Mendes, M. M. L.; Sinn, L.; O'Reilly, F. J.; Rappsilber, J. In Situ Structural Restraints from Cross-Linking Mass Spectrometry in Human Mitochondria. *J. Proteome Res.* **2020**, *19*, 327–336.

(848) Schweppe, D. K.; Chavez, J. D.; Lee, C. F.; Caudal, A.; Kruse, S. E.; Stuppard, R.; Marcinek, D. J.; Shadel, G. S.; Tian, R.; Bruce, J. E. Mitochondrial Protein Interactome Elucidated by Chemical Cross-Linking Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2017**, *114*, 1732–1737.

(849) Budayeva, H. G.; Kirkpatrick, D. S. Monitoring Protein Communities and Their Responses to Therapeutics. *Nat. Rev. Drug Discov.* **2020**, 1–13.

(850) Liu, Y.; Mi, Y.; Mueller, T.; Kreibich, S.; Williams, E. G.; Van Drogen, A.; Borel, C.; Frank, M.; Germain, P.-L.; Bludau, I.; et al. Multi-Omic Measurements of Heterogeneity in HeLa Cells across Laboratories. *Nat. Biotechnol.* **2019**, *37*, 314–322.

(851) Lund-Johansen, F.; Tran, T.; Mehta, A. Towards Reproducibility in Large-Scale Analysis of Protein–Protein Interactions. *Nat. Methods* **2021**, *18*, 720–721.

(852) Robertson, M. J.; Meyerowitz, J. G.; Skiniotis, G. Drug Discovery in the Era of Cryo-Electron Microscopy. *Trends Biochem. Sci.* **2022**, *47*, 124–135.

(853) Schaffer, M.; Pfeffer, S.; Mahamid, J.; Kleindiek, S.; Laugks, T.; Albert, S.; Engel, B. D.; Rummel, A.; Smith, A. J.; Baumeister, W.; et al. A Cryo-FIB Lift-out Technique Enables Molecular-Resolution Cryo-ET within Native Caenorhabditis Elegans Tissue. *Nat. Methods* **2019**, *16*, 757–762.

(854) Turk, M.; Baumeister, W. The Promise and the Challenges of Cryo-electron Tomography. *FEBS Lett.* **2020**, *594*, 3243–3261.

(855) Castaño-Díez, D.; Zanetti, G. In Situ Structure Determination by Subtomogram Averaging. *Curr. Opin. Struct. Biol.* **2019**, *58*, 68–75.

(856) Metskas, L. A.; Briggs, J. A. G. Fluorescence-Based Detection of Membrane Fusion State on a Cryo-EM Grid Using Correlated Cryo-Fluorescence and Cryo-Electron Microscopy. *Microsc. Microanal.* **2019**, *25*, 942–949.

(857) Li, X. Cryo-Electron Tomography: Observing the Cell at the Atomic Level. *Nat. Methods* **2021**, *18*, 440–441.

(858) Ronchi, P.; Mizzon, G.; Machado, P.; D'Imprima, E.; Best, B. T.; Cassella, L.; Schnorrenberg, S.; Montero, M. G.; Jechlinger, M.; Ephrussi, A.; et al. High-Precision Targeting Workflow for Volume Electron Microscopy. *J. Cell Biol.* **2021**, *220*, e202104069.

(859) Zila, V.; Margiotta, E.; Turoňová, B.; Müller, T. G.; Zimmerli, C. E.; Mattei, S.; Allegretti, M.; Börner, K.; Rada, J.; Müller, B.; et al. Cone-Shaped HIV-1 Capsids Are Transported through Intact Nuclear Pores. *Cell* **2021**, *184*, 1032-1046.e18.

(860) Vilas, J. L.; Oton, J.; Messaoudi, C.; Melero, R.; Conesa, P.; Ramirez-Aportela, E.; Mota, J.; Martinez, M.; Jimenez, A.; Marabini, R.; et al. Measurement of Local Resolution in Electron Tomography. *J. Struct. Biol. X* **2020**, *4*, 100016.

(861) Wagner, F. R.; Watanabe, R.; Schampers, R.; Singh, D.; Persoon, H.; Schaffer, M.; Fruhstorfer, P.; Plitzko, J.; Villa, E. Preparing Samples from Whole Cells Using Focused-Ion-Beam Milling for Cryo-Electron Tomography. *Nat. Protoc.* **2020**, 1–32.

(862) Guo, Q.; Lehmer, C.; Martínez-Sánchez, A.; Rudack, T.; Beck, F.; Hartmann, H.; Pérez-Berlanga, M.; Frottin, F.; Hipp, M. S.; Hartl, F. U.; et al. In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. *Cell* **2018**, *172*, 696-705.e12.

(863) Bäuerlein, F. J. B.; Fernández-Busnadiego, R.; Baumeister, W. Investigating the Structure of Neurotoxic Protein Aggregates Inside Cells. *Trends Cell Biol.* **2020**, *30*, 951–966.

(864) Trinkaus, V. A.; Riera-Tur, I.; Martínez-Sánchez, A.; Bäuerlein, F. J. B.; Guo, Q.; Arzberger, T.; Baumeister, W.; Dudanova, I.; Hipp, M. S.; Hartl, F. U.; et al. In Situ Architecture of Neuronal α-Synuclein Inclusions. *Nat. Commun.* **2021**, *12*, 2110.

(865) Watanabe, R.; Buschauer, R.; Böhning, J.; Audagnotto, M.; Lasker, K.; Lu, T.-W.; Boassa, D.; Taylor, S.; Villa, E. The In Situ Structure of Parkinson's Disease-Linked LRRK2. *Cell* **2020**, *182*, 1508-1518.e16.

(866) Fäßler, F.; Dimchev, G.; Hodirnau, V.-V.; Wan, W.; Schur, F. K. M. Cryo-Electron Tomography Structure of Arp2/3 Complex in Cells Reveals New Insights into the Branch Junction. *Nat. Commun.* **2020**, *11*, 6437.

(867) Burbaum, L.; Schneider, J.; Scholze, S.; Böttcher, R. T.; Baumeister, W.; Schwille, P.; Plitzko, J. M.; Jasnin, M. Molecular-Scale Visualization of Sarcomere Contraction within Native Cardiomyocytes. *Nat. Commun.* **2021**, *12*, 4086.

(868) Bykov, Y. S.; Schaffer, M.; Dodonova, S. O.; Albert, S.; Plitzko, J. M.; Baumeister, W.; Engel, B. D.; Briggs, J. A. The Structure of the COPI Coat Determined within the Cell. *eLife* **2017**, *6*, e32493.

(869) Rapisarda, C.; Cherrak, Y.; Kooger, R.; Schmidt, V.; Pellarin, R.; Logger, L.; Cascales, E.; Pilhofer, M.; Durand, E.; Fronzes, R. *In Situ* and High-resolution Cryo- EM Structure of a Bacterial Type VI Secretion System Membrane Complex. *EMBO J.* **2019**, *38*.

(870) Shi, X.; Chen, M.; Yu, Z.; Bell, J. M.; Wang, H.; Forrester, I.; Villarreal, H.; Jakana, J.; Du, D.; Luisi, B. F.; et al. In Situ Structure and Assembly of the Multidrug Efflux Pump AcrAB-TolC. *Nat. Commun.* **2019**, *10*, 2635.

(871) Ghosal, D.; Kim, K. W.; Zheng, H.; Kaplan, M.; Truchan, H. K.; Lopez, A. E.; McIntire, I. E.; Vogel, J. P.; Cianciotto, N. P.; Jensen, G. J. In Vivo Structure of the Legionella Type II Secretion System by Electron Cryotomography. *Nat. Microbiol.* **2019**, *4*, 2101–2108.

(872) Allegretti, M.; Zimmerli, C. E.; Rantos, V.; Wilfling, F.; Ronchi, P.; Fung, H. K. H.; Lee, C.-W.; Hagen, W.; Turoňová, B.; Karius, K.; et al. In-Cell Architecture of the Nuclear Pore and Snapshots of Its Turnover. *Nature* **2020**, *586*, 796–800.

(873) von Kügelgen, A.; Tang, H.; Hardy, G. G.; Kureisaite-Ciziene, D.; Brun, Y. V.; Stansfeld, P. J.; Robinson, C. V.; Bharat, T. A. M. In Situ Structure of an Intact Lipopolysaccharide-Bound Bacterial Surface Layer. *Cell* **2020**, *180*, 348-358.e15.

(874) Lucas, B. A.; Himes, B. A.; Xue, L.; Grant, T.; Mahamid, J.; Grigorieff, N. Locating Macromolecular Assemblies in Cells by 2D Template Matching with CisTEM. *eLife* **2021**, *10*, e68946.

(875) Turoňová, B.; Schur, F. K. M.; Wan, W.; Briggs, J. A. G. Efficient 3D-CTF Correction for Cryo-Electron Tomography Using NovaCTF Improves Subtomogram Averaging Resolution to 3.4 Å. *J. Struct. Biol.* **2017**, *199*, 187–195.

(876) Turoňová, B.; Hagen, W. J. H.; Obr, M.; Mosalaganti, S.; Beugelink, J. W.; Zimmerli, C. E.; Kräusslich, H.-G.; Beck, M. Benchmarking Tomographic Acquisition Schemes for High-Resolution Structural Biology. *Nat. Commun.* **2020**, *11*, 876.

(877) Bouvette, J.; Liu, H.-F.; Du, X.; Zhou, Y.; Sikkema, A. P.; da Fonseca Rezende e Mello, J.; Klemm, B. P.; Huang, R.; Schaaper, R. M.; Borgnia, M. J.; et al. Beam Image-Shift Accelerated Data Acquisition for near-Atomic Resolution Single-Particle Cryo-Electron Tomography. *Nat. Commun.* **2021**, *12*, 1957.

(878) Tegunov, D.; Xue, L.; Dienemann, C.; Cramer, P.; Mahamid, J. Multi-Particle Cryo-EM Refinement with M Visualizes Ribosome-Antibiotic Complex at 3.5 Å in Cells. *Nat. Methods* **2021**, *18*, 186–193.

(879) Yeates, T. O.; Agdanowski, M. P.; Liu, Y. Development of Imaging Scaffolds for Cryo-Electron Microscopy. *Curr. Opin. Struct. Biol.* **2020**, *60*, 142–149.

(880) Uchański, T.; Masiulis, S.; Fischer, B.; Kalichuk, V.; López-Sánchez, U.; Zarkadas, E.; Weckener, M.; Sente, A.; Ward, P.; Wohlkönig, A.; et al. Megabodies Expand the Nanobody Toolkit

for Protein Structure Determination by Single-Particle Cryo-EM. *Nat. Methods* **2021**, *18*, 60–68.

(881) Kyrilis, F. L.; Meister, A.; Kastritis, P. L. Integrative Biology of Native Cell Extracts: A New Era for Structural Characterization of Life Processes. *Biol. Chem.* **2019**, *400*, 831–846.

(882) Weissenberger, G.; Henderikx, R. J. M.; Peters, P. J. Understanding the Invisible Hands of Sample Preparation for Cryo-EM. *Nat. Methods* **2021**, *18*, 463–471.

(883) Pyle, E.; Zanetti, G. Current Data Processing Strategies for Cryo-Electron Tomography and Subtomogram Averaging. *Biochem. J.* **2021**, *478*, 1827–1845.

(884) Klebl, D. P.; Gravett, M. S. C.; Kontziampasis, D.; Wright, D. J.; Bon, R. S.; Monteiro, D. C. F.; Trebbin, M.; Sobott, F.; White, H. D.; Darrow, M. C.; et al. Need for Speed: Examining Protein Behavior during CryoEM Grid Preparation at Different Timescales. *Struct. Des.* **2020**, *28*, 1238-1248.e4.

(885) Bradford, S. Y. C.; El Khoury, L.; Ge, Y.; Osato, M.; Mobley, D. L.; Fischer, M. Temperature Artifacts in Protein Structures Bias Ligand-Binding Predictions. *Chem. Sci.* **2021**, *12*, 11275–11293.

(886) Pratapa, A.; Doron, M.; Caicedo, J. C. Image-Based Cell Phenotyping with Deep Learning. *Curr. Opin. Chem. Biol.* **2021**, *65*, 9–17.

(887) Singh, H.; Tiwari, K.; Tiwari, R.; Pramanik, S. K.; Das, A. Small Molecule as Fluorescent Probes for Monitoring Intracellular Enzymatic Transformations. *Chem. Rev.* **2019**, *119*, 11718–11760.

(888) Olsen, R. H. J.; DiBerto, J. F.; English, J. G.; Glaudin, A. M.; Krumm, B. E.; Slocum, S. T.; Che, T.; Gavin, A. C.; McCorvy, J. D.; Roth, B. L.; et al. TRUPATH, an Open-Source Biosensor Platform for Interrogating the GPCR Transducerome. *Nat. Chem. Biol.* **2020**, 1–13.

(889) Herholt, A.; Galinski, S.; Geyer, P. E.; Rossner, M. J.; Wehr, M. C. Multiparametric Assays for Accelerating Early Drug Discovery. *Trends Pharmacol. Sci.* **2020**, *41*, 318–335.

(890) Gorgulla, C.; Boeszoermenyi, A.; Wang, Z.-F.; Fischer, P. D.; Coote, P. W.; Padmanabha Das, K. M.; Malets, Y. S.; Radchenko, D. S.; Moroz, Y. S.; Scott, D. A.; et al. An Open-Source Drug Discovery Platform Enables Ultra-Large Virtual Screens. *Nature* **2020**, *580*, 663–668.

(891) Ghislat, G.; Rahman, T.; Ballester, P. J. Recent Progress on the Prospective Application of Machine Learning to Structure-Based Virtual Screening. *Curr. Opin. Chem. Biol.* **2021**, *65*, 28–34.

(892) Carvalho, J.; Alves, S.; Castro, M. M. C. A.; Geraldes, C. F. G. C.; Queiroz, J. A.; Fonseca, C. P.; Cruz, C. Development of a Bioreactor System for Cytotoxic Evaluation of Pharmacological Compounds in Living Cells Using NMR Spectroscopy. *J. Pharmacol. Toxicol. Methods* **2019**, *95*, 70–78.

(893) Hertig, D.; Maddah, S.; Memedovski, R.; Kurth, S.; Moreno, A.; Pennestri, M.; Felser, A.; Nuoffer, J.-M.; Vermathen, P. Live Monitoring of Cellular Metabolism and Mitochondrial Respiration in 3D Cell Culture System Using NMR Spectroscopy. *Analyst* **2021**, *146*, 4326–4339.

(894) Knitsch, R.; AlWahsh, M.; Raschke, H.; Lambert, J.; Hergenröder, R. In Vitro Spatio-Temporal NMR Metabolomics of Living 3D Cell Models. *Anal. Chem.* **2021**, acs.analchem.1c02221.

(895) d'Avignon, D. A.; Ge, X. In Vivo NMR Investigations of Glyphosate Influences on Plant Metabolism. *J. Magn. Reson. San Diego Calif 1997* **2018**, *292*, 59–72.

(896) Nguyen, T. T. M.; An, Y. J.; Cha, J. W.; Ko, Y.-J.; Lee, H.; Chung, C. H.; Jeon, S.-M.; Lee, J.; Park, S. Real-Time In-Organism NMR Metabolomics Reveals Different Roles of AMP-Activated Protein Kinase Catalytic Subunits. *Anal. Chem.* **2020**, acs.analchem.9b05670.

(897) Lane, D.; Bermel, W.; Ning, P.; Jeong, T.-Y.; Martin, R.; Soong, R.; Wu, B.; Tabatabaei-Anaraki, M.; Heumann, H.; Gundy, M.; et al. Targeting the Lowest Concentration of a Toxin That Induces a Detectable Metabolic Response in Living Organisms: Time Resolved In Vivo 2D NMR During a Concentration Ramp. *Anal. Chem.* **2020**, acs.analchem.0c01370-14.

(898) Tabatabaei-Anaraki, M.; Dutta Majumdar, R.; Wagner, N.; Soong, R.; Kovacevic, V.; Reiner, E. J.; Bhavsar, S. P.; Ortiz Almirall, X.; Lane, D.; Simpson, M. J.; et al. Development and Application of a Low-Volume Flow System for Solution-State in Vivo NMR. *Anal. Chem.* **2018**, *90*, 7912–7921.

(899) Fugariu, I.; Bermel, W.; Lane, D.; Soong, R.; Simpson, A. J. In-Phase Ultra High-Resolution In Vivo NMR. *Angew. Chem.* **2017**, *129*, 6421–6425.

(900) Lane, D.; Skinner, T. E.; Gershenzon, N. I.; Bermel, W.; Soong, R.; Dutta Majumdar, R.; Liaghati-Mobarhan, Y.; Schmidt, S.; Heumann, H.; Monette, M.; et al. Assessing the Potential of Quantitative 2D HSQC NMR in 13C Enriched Living Organisms. *J. Biomol. NMR* **2019**, *73*, 31–42.

(901) Zhan, H.; Huang, Y.; Chen, Z. High-Resolution Probing of Heterogeneous Samples by Spatially Selective Pure Shift NMR Spectroscopy. *J. Phys. Chem. Lett.* **2019**, *10*, 7356–7361.

(902) Thankamony, A. S. L.; Wittmann, J. J.; Kaushik, M.; Corzilius, B. Dynamic Nuclear Polarization for Sensitivity Enhancement in Modern Solid-State NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *102–103*, 120–195.

(903) Plainchont, B.; Berruyer, P.; Dumez, J.-N.; Jannin, S.; Giraudeau, P. Dynamic Nuclear Polarization Opens New Perspectives for NMR Spectroscopy in Analytical Chemistry. *Anal. Chem.* **2018**, *90*, 3639–3650.

(904) Pinon, A. C.; Capozzi, A.; Ardenkjaer-Larsen, J. H. Hyperpolarization via Dissolution Dynamic Nuclear Polarization: New Technological and Methodological Advances. *Magn. Reson. Mater. Phys. Biol. Med.* **2020**, *137*, 8428–19.

(905) Hövener, J.-B.; Pravdivtsev, A. N.; Kidd, B.; Bowers, C. R.; Glöggler, S.; Kovtunov, K. V.; Plaumann, M.; Katz-Brull, R.; Buckenmaier, K.; Jerschow, A.; et al. Parahydrogen-Based Hyperpolarization for Biomedicine. *Angew. Chem. Int. Ed.* **2018**, *57*, 11140–11162.

(906) Barskiy, D. A.; Knecht, S.; Yurkovskaya, A. V.; Ivanov, K. L. SABRE: Chemical Kinetics and Spin Dynamics of the Formation of Hyperpolarization. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *114–115*, 33–70.

(907) Kondo, Y.; Nonaka, H.; Takakusagi, Y.; Sando, S. Design of Nuclear Magnetic Resonance Molecular Probes for Hyperpolarized Bioimaging. *Angew. Chem. Int. Ed.* **2021**, anie.201915718.

(908) Schmidt, A. B.; Wörner, J.; Pravdivtsev, A.; Knecht, S.; Scherer, H.; Weber, S.; Hennig, J.; Elverfeldt, D.; Hövener, J. B. Lifetime of Parahydrogen in Aqueous Solutions and Human Blood. *ChemPhysChem* **2019**, *20*, 2408–2412.

(909) Berthault, P.; Boutin, C.; Martineau-Corcos, C.; Carret, G. Use of Dissolved Hyperpolarized Species in NMR: Practical Considerations. *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *118–119*, 74–90.

(910) Reineri, F.; Cavallari, E.; Carrera, C.; Aime, S. Hydrogenative-PHIP Polarized Metabolites for Biological Studies. *Magn. Reson. Mater. Phys. Biol. Med.* **2021**, *34*, 25–47.

(911) Narasimhan, S.; Folkers, G. E.; Baldus, M. When Small Becomes Too Big: Expanding the Use of In-Cell Solid-State NMR Spectroscopy. *ChemPlusChem* **2020**, *85*, 760–768.

(912) Jaudzems, K.; Polenova, T.; Pintacuda, G.; Oschkinat, H.; Lesage, A. DNP NMR of Biomolecular Assemblies. *J. Struct. Biol.* **2019**, *206*, 90–98.

(913) Cai, S.; Seu, C.; Kovacs, Z.; Sherry, A. D.; Chen, Y. Sensitivity Enhancement of Multidimensional NMR Experiments by Paramagnetic Relaxation Effects. *J. Am. Chem. Soc.* **2006**, *128*, 13474–13478.

(914) Mulder, F. A. A.; Tenori, L.; Luchinat, C. Fast and Quantitative NMR Metabolite Analysis Afforded by a Paramagnetic Co-Solute. *Angew. Chem. Int. Ed.* **2019**, *58*, 15283–15286.

(915) Theillet, F.-X.; Binolfi, A.; Liokatis, S.; Verzini, S.; Selenko, P. Paramagnetic Relaxation Enhancement to Improve Sensitivity of Fast NMR Methods: Application to Intrinsically Disordered Proteins. *J. Biomol. NMR* **2011**, *51*, 487–495.

(916) Mayzel, M.; Rosenlöw, J.; Isaksson, L.; Orekhov, V. Y. Time-Resolved Multidimensional NMR with Non-Uniform Sampling. *J. Biomol. NMR* **2014**, *58*, 129–139.

(917) Dass, R.; Grudziąż, K.; Ishikawa, T.; Nowakowski, M.; Dębowska, R.; Kazimierczuk, K. Fast 2D NMR Spectroscopy for In Vivo Monitoring of Bacterial Metabolism in Complex Mixtures. *Front. Microbiol.* **2017**, *8*, 968–12.

(918) Gołowicz, D.; Kasprzak, P.; Orekhov, V.; Kazimierczuk, K. Fast Time-Resolved NMR with Non-Uniform Sampling. *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *116*, 40–55.

(919) Urbańczyk, M.; Shchukina, A.; Gołowicz, D.; Kazimierczuk, K. TReNDS-Software for Reaction Monitoring with Time-Resolved Non-Uniform Sampling: TReNDS - Software for Reaction Monitoring with Time-Resolved Non-Uniform Sampling. *Magn. Reson. Chem.* **2019**, *57*, 4–12.

(920) Qu, X.; Huang, Y.; Lu, H.; Qiu, T.; Guo, D.; Agback, T.; Orekhov, V.; Chen, Z. Accelerated Nuclear Magnetic Resonance Spectroscopy with Deep Learning. *Angew. Chem.* **2020**, *59*, 10297–10300.

(921) Karunanithy, G.; Hansen, D. F. FID-Net: A Versatile Deep Neural Network Architecture for NMR Spectral Reconstruction and Virtual Decoupling. *J. Biomol. NMR* **2021**, *75*, 179–191.

(922) Karunanithy, G.; Mackenzie, H. W.; Hansen, D. F. Virtual Homonuclear Decoupling in Direct Detection Nuclear Magnetic Resonance Experiments Using Deep Neural Networks. *J. Am. Chem. Soc.* **2021**, jacs.1c04010.

(923) Man, P. P.; Bonhomme, C.; Babonneau, F. Denoising NMR Time-Domain Signal by Singular-Value Decomposition Accelerated by Graphics Processing Units. *Solid State Nucl. Magn. Reson.* **2014**, *61–62*, 28–34.

(924) Bruno, F.; Francischello, R.; Bellomo, G.; Gigli, L.; Flori, A.; Menichetti, L.; Tenori, L.; Luchinat, C.; Ravera, E. Multivariate Curve Resolution for 2D Solid-State NMR Spectra. *Anal. Chem.* **2020**, *92*, 4451–4458.

(925) Mobarhan, Y. L.; Fortier-McGill, B.; Soong, R.; Maas, W. E.; Fey, M.; Monette, M.; Stronks, H. J.; Schmidt, S.; Heumann, H.; Norwood, W.; et al. Comprehensive Multiphase NMR Applied to a Living Organism. *Chem Sci* **2016**, *7*, 4856–4866.

(926) Li, D.-W.; Hansen, A. L.; Yuan, C.; Bruschweiler-Li, L.; Brüschweiler, R. DEEP Picker Is a Deep Neural Network for Accurate Deconvolution of Complex Two-Dimensional NMR Spectra. *Nat. Commun.* **2021**, *12*, 5229.

(927) Zweckstetter, M. NMR Hawk-eyed View of AlphaFold2 Structures. *Protein Sci.* **2021**, *30*, 2333–2337.

(928) Huang, Y. J.; Zhang, N.; Bersch, B.; Fidelis, K.; Inouye, M.; Ishida, Y.; Kryshtafovych, A.; Kobayashi, N.; Kuroda, Y.; Liu, G.; et al. Assessment of Prediction Methods for Protein Structures Determined by NMR in CASP14 : Impact of AlphaFold2. *Proteins Struct. Funct. Bioinforma.* **2021**, prot.26246.

(929) Müller, D.; Trucks, S.; Schwalbe, H.; Hengesbach, M. Genetic Code Expansion Facilitates Position-Selective Modification of Nucleic Acids and Proteins. *ChemPlusChem* **2020**, *85*, 1233–1243.

(930) Wang, X.; Liu, D.; Shen, L.; Li, F.; Li, Y.; Yang, L.; Xu, T.; Tao, H.; Yao, D.; Wu, L.; et al. A Genetically Encoded F-19 NMR Probe Reveals the Allosteric Modulation Mechanism of Cannabinoid Receptor 1. *J. Am. Chem. Soc.* **2021**, *143*, 16320–16325.

(931) Walstein, K.; Petrovic, A.; Pan, D.; Hagemeier, B.; Vogt, D.; Vetter, I. R.; Musacchio, A. Assembly Principles and Stoichiometry of a Complete Human Kinetochore Module. *Sci. Adv.* **2021**, *7*, eabg1037.

(932) Haller, J. D.; Bodor, A.; Luy, B. Real-Time Pure Shift Measurements for Uniformly Isotope-Labeled Molecules Using X-Selective BIRD Homonuclear Decoupling. *J. Magn. Reson.* **2019**, *302*, 64–71.

(933) Bodor, A.; Haller, J. D.; Bouguechtouli, C.; Theillet, F.-X.; Nyitray, L.; Luy, B. Power of Pure Shift HαCα Correlations: A Way to Characterize Biomolecules under Physiological Conditions. *Anal. Chem.* **2020**, *92*, 12423–12428.

(934) Wong, L. E.; Kim, T. H.; Muhandiram, D. R.; Forman-Kay, J. D.; Kay, L. E. NMR Experiments for Studies of Dilute and Condensed Protein Phases: Application to the Phase-Separating Protein CAPRIN1. *J. Am. Chem. Soc.* **2020**, *142*, 2471–2489.

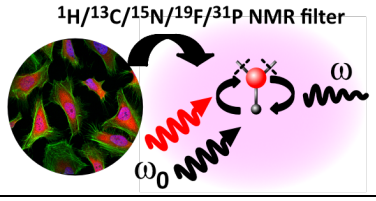
(935) Takeuchi, K.; Arthanari, H.; Shimada, I.; Wagner, G. Nitrogen Detected TROSY at High Field Yields High Resolution and Sensitivity for Protein NMR. *J. Biomol. NMR* **2015**, *63*, 323–331.

(936) Chhabra, S.; Fischer, P.; Takeuchi, K.; Dubey, A.; Ziarek, J. J.; Boeszoermenyi, A.; Mathieu, D.; Bermel, W.; Davey, N. E.; Wagner, G.; et al. 15N Detection Harnesses the Slow Relaxation Property of Nitrogen: Delivering Enhanced Resolution for Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115*, E1710–E1719.

(937) Tokunaga, Y.; Takeuchi, K.; Okude, J.; Ori, K.; Torizawa, T.; Shimada, I. Structural Fingerprints of an Intact Monoclonal Antibody Acquired under Formulated Storage Conditions via ^{15}N Direct Detection Nuclear Magnetic Resonance. *J. Med. Chem.* **2020**, *63*, 5360–5366.

(938) Karunanithy, G.; Shukla, V. K.; Hansen, D. F. Methodological Advancements for Characterising Protein Side Chains by NMR Spectroscopy. *Curr. Opin. Struct. Biol.* **2021**, *70*, 61–69.

(939) Bastawrous, M.; Tabatabaei-Anaraki, M.; Soong, R.; Bermel, W.; Gundy, M.; Boenisch, H.; Heumann, H.; Simpson, A. J. Inverse or Direct Detect Experiments and Probes: Which Are "Best" for in-Vivo NMR Research of 13C Enriched Organisms? *Anal. Chim. Acta* **2020**, *1138*, 168–180.


$^{1}H/^{13}C/^{15}N/^{19}F/^{31}P$ NMR filter
ω
$ω_0$