Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 158 (2007) 244–253
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Review

Looking into live cells with in-cell NMR spectroscopy

Philipp Selenko *, Gerhard Wagner

Harvard Medical School, Department of Biological Chemistry and Molecular Pharmacology (BCMP), 240 Longwood Avenue, Boston, MA 02115, USA

Received 24 May 2006; received in revised form 2 April 2007; accepted 3 April 2007
Available online 10 April 2007

Abstract

In-cell NMR spectroscopy has gained recent popularity since it provides means to analyze the conformational and functional properties of proteins inside living cells and at atomic resolution. High-resolution in-cell NMR spectroscopy was originally established in bacterial cells and based on a rationale that relies on protein over-expression and sample analysis within the same cellular environment. Here, we review in-cell NMR approaches in *Xenopus laevis* oocytes and evaluate potential future applications in other eukaryotic cell types.

Keywords: High-resolution NMR spectroscopy; In-cell NMR spectroscopy; Protein transduction; *Xenopus laevis* oocytes

1. High-resolution in-cell NMR spectroscopy

Biophysical methods for the structural characterization of biomolecules are often confined to artificial, *in vitro* experimental setups. X-ray crystallography and high-resolution electron microscopy are intrinsically restricted from *in vivo* approaches due to their requirement for pure samples and crystalline or vitrified specimens. Nuclear magnetic resonance (NMR) spectroscopy, the only other method for structural investigations at the atomic level, allows for the direct observation of NMR-active nuclei within any NMR-inactive environment and can thus be employed to analyze biomolecules *in vivo* and inside cells (Serber et al., 2005). Historically, small molecule *in vivo* NMR spectroscopy denotes the observation of metabolites in suspensions of bacteria and other cells by means of investigating a few characteristic proton or phosphor resonance signals in 1-dimensional NMR spectra (Cohen et al., 1989; Szwergold, 1992). In-cell NMR spectroscopy, in contrast, employs modern methods of isotope labeling and multi-dimensional, isotope-edited correlation experiments to obtain structural information on proteins within living cells (Reckel et al., 2005). In brief, this method enables high-resolution snapshots of intracellular protein conformations. These structural 'fingerprints' may change upon biological interactions, post-translational protein modifications or due to structural rearrangements. In-cell NMR measurements then detect these biological events and read out residue specific changes in a time dependent- and quantitative manner. Why would these measurements be of any additional value compared to conventional analyses by *in vitro* methods? First, most proteins function inside cells and in a highly crowded, viscous solution that harbors an intricate network of biological activities simultaneously exerted by a large number of macromolecules. Whereas *in vitro* structural analyses on pure samples have shaped our 3-dimensional understanding of many biological processes, they do not necessarily reflect the true nature of the cellular environment. What are the biological questions that can be addressed by in-cell NMR techniques? Does an *in vitro* determined protein structure represent the cellular *in vivo* conformation? How do proteins that do not exhibit folded properties in their pure states behave in a cellular environment? How do *in vitro* investigated conformational changes, upon ligand binding for example, relate to equivalent structural alterations *in vivo*? How do post-translational protein modifications affect protein structure and in what ways does a protein respond structurally to cellular processes like apoptosis, cell cycle progression or

* Corresponding author.
E-mail address: philipp_selenko@hms.harvard.edu (P. Selenko).

1047-8477/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2007.04.001

differentiation? Clearly, these are important issues that can only be adequately addressed in a cellular setting and by a high-resolution method. Second, biological reactions often involve the dynamic modulation of a protein's activity. In eukaryotes, this is often achieved by reversible post-translational modifications of one, or many amino acid residues in a protein. Such modifications involve the exchange of various functional groups in highly regulated processes that are typically controlled by intrinsic cellular mechanisms. Moreover, these alterations often result in substantial conformational rearrangements, which lead to specific modulations in a protein's function or activity. Eukaryotic in-cell NMR techniques in particular enable the direct observation of these cellular reactions, and of their structural implications, and can thus provide important new insights into the biological behavior of a protein.

In the following chapters we will introduce the basic concepts of NMR experiments inside living cells, provide a brief introduction to NMR properties of labeled proteins in a cellular environment, review in-cell NMR approaches in bacterial and eukaryotic cells and outline future in-cell NMR applications.

2. Basic concepts

Today, many in-cell NMR applications employ bacterial cells and follow the experimental rationale proposed by Serber et al. (2001). Their approach exploits the fact that most atomic nuclei in natural substances are NMR-inactive and hence not directly observable by NMR spectroscopy. Modern NMR methods utilize recombinant protein over-expression and isotope labeling to substitute some of these inactive nuclei (^{14}N, ^{12}C) with NMR-active isotopes (^{15}N, ^{13}C). Multi-dimensional correlation experiments then allow the detection of pairs of coupled nuclei in labeled proteins, such as ^{1}H and ^{15}N or ^{1}H and ^{13}C, in any aqueous environments that do not contain significant amounts of such isotope pairs. In essence, isotope labeling functions as a selective filter that renders the NMR-inactive cellular environment invisible to spectroscopic eye. In that sense, the underlying principle of in-cell NMR spectroscopy is very similar to applications in cellular microscopy that employ fluorescent labeled proteins.

Isotope labeling involves recombinant protein production in growth media that provide isotope-substituted metabolic precursors (^{15}N-ammonium chloride, ^{13}C-glucose). Due to the high protein expression levels that are typically achieved with viral promoters and polymerases in *Escherichia coli,* recombinant polypeptides accumulate rapidly post-induction and generally outperform protein synthesis rates of endogenous *E. coli* gene products. When, in addition, bacterial cells are grown in unlabeled medium first and only switched to labeled growth conditions before the induction of recombinant protein expression, selective isotope labeling is restricted to the recombinant protein only. Hence, and without further purification, isotope-edited correlation experiments on intact cells yield high-resolution NMR spectra of the intracellular recombinant protein (Fig. 1a). Alternatively, labeled protein samples can be conventionally purified from bacteria and then transferred into other cells for in-cell NMR analyses (Fig. 1b). For eukaryotic in-cell NMR applications in *Xenopus laevis* oocytes, for example, labeled proteins are deposited by microinjection. It is evident that the same underlying principle constitutes the experimental feasibility of both of these methods.

3. NMR parameters for in-cell NMR experiments

In the following paragraph we will briefly outline important NMR parameters that are routinely employed for the qualitative and quantitative interpretation of in-cell, and other, NMR spectra.

Two-dimensional correlation experiments of ^{1}H and ^{15}N, or ^{13}C, serve as the primary in-cell NMR techniques. The correlation of these NMR-active atomic nuclei, by means of specifically tailored NMR pulse-sequences, yields individual NMR signals (or resonance cross-peaks) at the respective resonance frequencies (or chemical shift values, $\delta[^{1}\mathrm{H}]$, $\delta[^{15}\mathrm{N}]$ and/or $\delta[^{13}\mathrm{C}]$) (Fig. 2a). For a folded protein, the resonance frequencies of individual residues are determined by the amino acid specific chemical environments, which in turn are defined by the protein's 3-dimensional structure. This leads to characteristic patterns of resonance cross-peaks in 2-dimensional correlation spectra, which reflect the overall conformational state of the labeled protein. Local changes in the chemical environment of labeled residues, either by ligand binding or conformational rearrangements, result in differences in resonance frequencies and changes in chemical shift values (Fig. 2b). The overall number of resonance peaks remains the same because both reactions do not introduce additional observable spin-pairs. Differences in resonance frequencies of participating residues are measured as chemical shift changes ($\Delta\delta_{\mathrm{total}} = \Delta\delta[^{1}\mathrm{H}] + \Delta\delta[^{15}\mathrm{N}$ or $^{13}\mathrm{C}]$) and mapped onto the 3-dimensional protein structure. Chemical shift changes serve as unique indicators of conformational alterations or localized binding events both *in vitro* and *in vivo* (Zuiderweg, 2002).

Second, the characteristic appearance of each NMR signal contains additional information about the spin system under investigation (Fig. 3). A prerequisite for liquid state NMR spectroscopy is that the molecule of interest tumbles freely in solution. The resulting overall tumbling rate depends on the size of the molecule and the temperature and viscosity of the NMR sample solution. These parameters determine the overall line widths of the respective NMR resonance signals. Yet, NMR resonances from one protein do not all exhibit identical peak intensities or uniform NMR line widths. Additional parameters like internal residue mobility, or solvent and conformational exchange, differentially affect the relaxation properties of individual spins and hence the appearance of the respective NMR signals (Fischer et al., 1998; Palmer, 2001; Peng and Wagner, 1994). Amino acids of

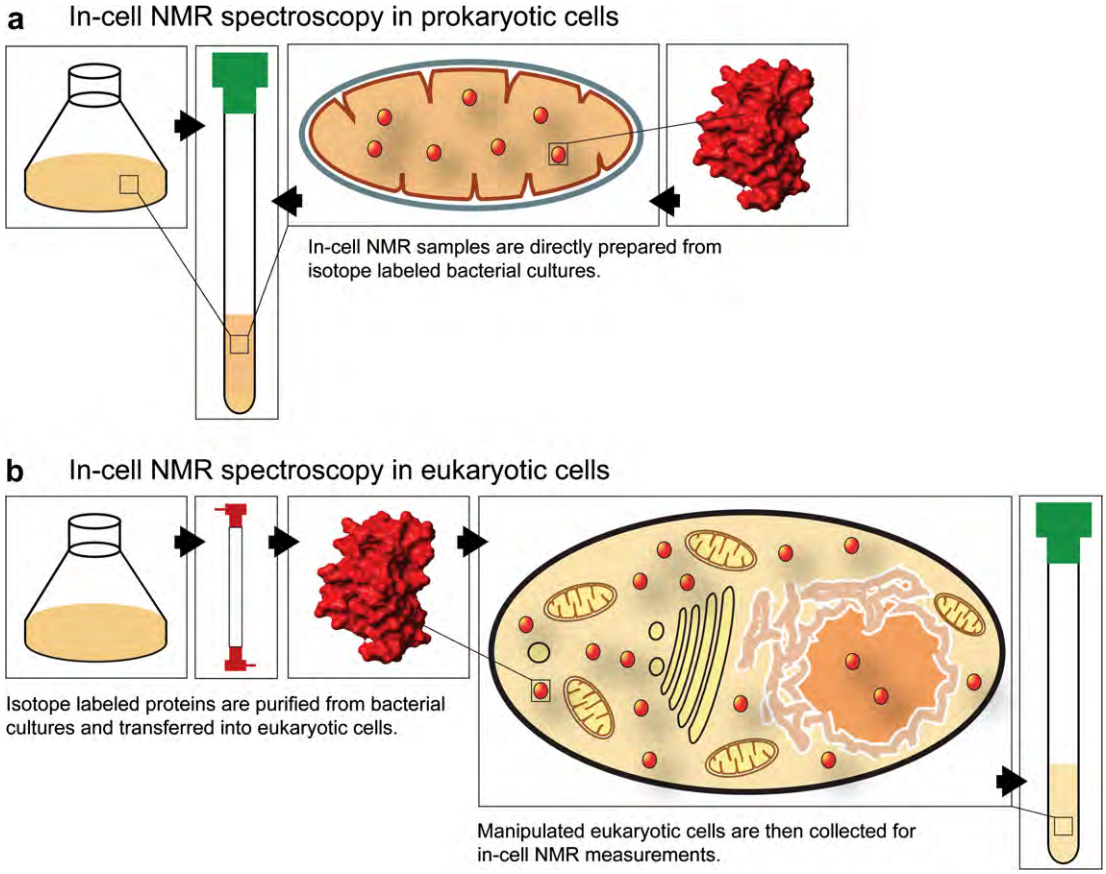

Fig. 1. (a) Prokaryotic in-cell NMR approaches typically employ recombinant protein over-expression, isotope labeling and in-cell NMR analyses within the same cell type. Suspensions of bacterial cells are directly analyzed without purification of the recombinant protein. (b) Eukaryotic in-cell NMR applications can involve isotope labeling in bacterial cells and recombinant protein purification prior to in-cell NMR sample preparation. Labeled proteins are then transferred into eukaryotic cells by microinjection, or other vector-based transduction techniques.

unstructured loop regions typically display narrower and more intense resonance signals than residues in secondary structure elements. These dynamic relaxation properties may undergo differential alterations in a cellular environment and the comparative analysis of changes in NMR line widths can therefore provide information about *in vivo* dynamics and exchange behaviors. In general, small proteins display large tumbling rates, which lead to slow overall relaxation and narrow NMR line widths (Fig. 3a). Molecules of larger size tumble more slowly, relax faster and exhibit broader resonance signals (Fig. 3b). Because the overall rotational tumbling rate is a direct function of the viscosity of the medium in which the macromolecule is dissolved, intracellular viscosity becomes a crucial parameter for in-cell NMR experiments. Any molecule in a cellular environment exhibits a reduced tumbling rate due to intracellular viscosity and hence displays broader NMR line widths (Fig. 3c). In the absence of protein binding to endogenous cellular factors, a direct comparison of individual protein line widths in buffer and in in-cell NMR experiments will readily yield a qualitative estimate about intracellular viscosity. Upon sample binding to cellular components, the resulting protein complexes can either display tumbling rates that correspond to the sum of their individual masses (Fig. 3d), or individual contributions to a mixed set of rates, when the interaction is restricted to a subset of residues (Fig. 3e). The latter results in residue specific line broadening, which yields information on the dynamics and localization of the cellular interaction. Binding to quasi-static cellular structures like organelles or membranes results in severe line broadening (Fig. 3f), which can serve as a qualitative indicator for the kind of interaction. Many biological binding events are dynamic and modulated by cellular signaling, which often leads to transient and interpretable changes of NMR line widths. It is apparent that a biologically active protein can experience any of the aforementioned conditions, and superpositions thereof, in a cellular environment. Complicated or poor quality in-cell NMR spectra are the likely result. In such cases, the researcher needs to reduce the complexity of the system under investigation, which can be achieved either by 'chopping up' full-length proteins into individual domains in order to selectively probe differential biological activities or by introducing site-directed mutations that abolish certain functional characteristics and similarly enable one to discriminate between specific cellular contributions.

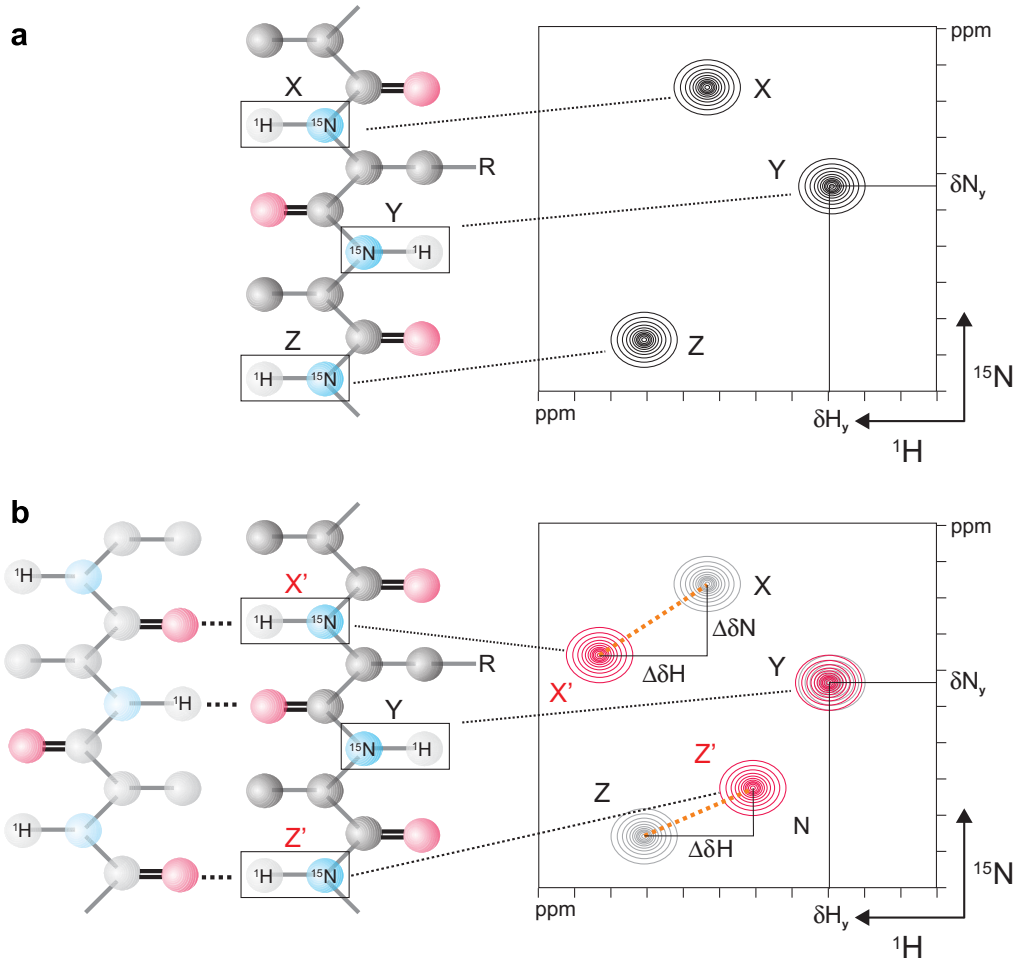

Fig. 2. (a) Schematic representation of a 2-dimensional ^{1}H, ^{15}N isotope-edited NMR spectrum. Backbone cross-peaks of amide spin-pairs give rise to a characteristic pattern of NMR resonance signals. The chemical shift (δ) of these peaks depends on the environment of the labeled spin-pairs, which is defined by their structural context. (b) Upon ligand binding, a limited set of amide spin-pairs experience a novel chemical environment, which results in selective chemical shift changes ($\Delta\delta$). Since those changes are more pronounced for residues directly involved in binding, their mapping onto the 3-dimensional structure of the protein identifies the interaction surface.

In summary, a thorough and comparative analysis of NMR chemical shift values and NMR line widths of labeled protein samples in different *in vitro* and *in vivo* environments can yield a wealth of structural and dynamic information about a protein's cellular behavior and about its potential interactions with endogenous cellular components.

4. In-cell NMR spectroscopy in prokaryotic cells

Several in-cell NMR applications for structural and functional studies of proteins in bacteria have been reported (Dedmon et al., 2002; Hubbard et al., 2003; McNulty et al., 2006; Serber et al., 2004; Wieruszeski et al., 2001). Bacterial in-cell NMR techniques have been successfully employed to analyze protein dynamics (Bryant, 2006; Bryant et al., 2005), protein–protein interactions (Burz et al., 2006), for de-novo resonance assignments (Reardon and Spicer, 2005) and automated structure determinations in crude cell extracts (Etezady-Esfarjani et al., 2006). Some of these prokaryotic in-cell NMR applications have been excellently reviewed in recent publications (Reckel et al., 2005; Serber et al., 2005). We will not further elaborate on these bacterial studies but restrict ourselves, for the remainder of the manuscript, to outline the methodological considerations for in-cell NMR applications in eukaryotic cells.

5. In-cell NMR spectroscopy in eukaryotic cells

Why do we wish to extend the applicability of in-cell NMR measurements to eukaryotic cells? Above all, prokaryotic organisms exhibit a limited range of biological activities and many of the cellular processes that define the prevalent topics in modern biological research are absent in bacteria. Post-translational protein modifications, for example, serve to ubiquitously regulate biological activities in eukaryotes, but are much less common in prokaryotic organisms. The presence of organelles and the

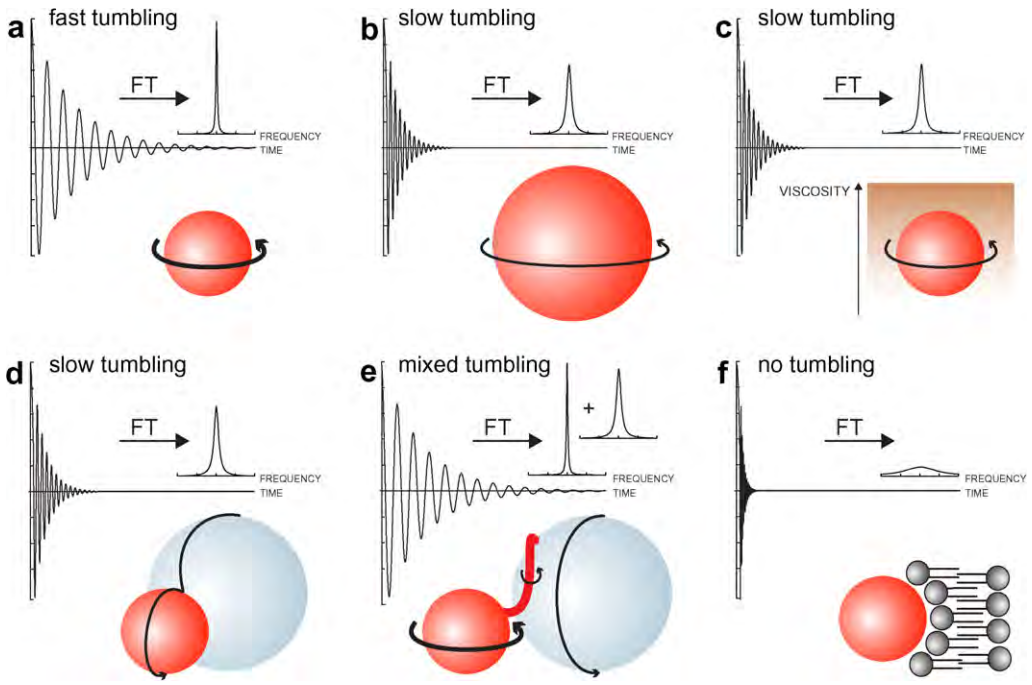

Fig. 3. (a) Small molecules display large rates of molecular tumbling, indicated by a thick arrow. This results in a slow decay of the experimentally observed time domain NMR signal (the free induction decay, or FID). Upon Fourier transformation (FT) of the time-domain FID signal, the frequency-domain NMR signal displays the characteristic Lorentzian line shape. (b) Large molecules tumble more slowly, indicated by a thin arrow, which results in fast relaxation rates and broad NMR line widths with reduced overall signal intensities. (c) An increase in viscosity of the sample solution results in an overall reduction in tumbling rates. Under these circumstances small molecules display relaxation properties, and hence NMR line widths, comparable to those of larger molecules in low viscosity environments. (d) Intermolecular interactions with components of larger size yield relaxation rates that correspond to the overall molecular weight of the complex (typically the sum of the individual rates). (e) Interactions with subsets of residues of the labeled molecule result in a mixed set of relaxation rates. These may yield non-uniform degrees of line broadening. (f) Interactions with immobile cellular structures, like membranes, effectively abolish molecular tumbling of the labeled molecule. Relaxation rates become so large that the resulting frequency domain signals are virtually undetectable.

concomitant requirement of regulated cellular transport constitutes another characteristic of eukaryotic cell identity. Compartmentalization *per se* results in the creation of sub-cellular environments with different physical and biological properties and little is known about the effects that these compartments exert on a protein's structure or function. Moreover, eukaryotic organisms display a high degree of cellular differentiation, which leads to functional specification and differences in biological activities in neighboring cell types. Do these specialized cellular environments differentially affect protein structure? Some protein examples indeed suggest so (Dyson and Wright, 2005). Together, we believe that eukaryotic model systems for in-cell NMR measurements will prove instrumental in structurally addressing a multitude of fundamental biological activities that are present in higher organisms.

What constitutes a suitable eukaryotic in-cell NMR system? First, we have to decide on the rationale for introducing labeled protein samples into the native, intracellular environment. We could either choose the original approach of sample over-expression and analysis within the same cell type, or turn towards intracellular sample delivery by alternative methods. Whereas recombinant protein over-expression in yeast-, insect- or mammalian cells would represent a 'eukaryotic solution' along the originally proposed rationale, the manipulation of *X. laevis* oocytes by microinjection represents a powerful alternative approach. Before we outline the experimental conditions for both methods, we will address some of the conceptual differences behind these two techniques.

Protein over-expression and in-cell NMR measurements within the same cell type do not require purification of the labeled polypeptide, which can be beneficial when a protein is difficult to isolate or complicated to stabilize in its pure form. On the other hand, in-cell protein production cannot be easily quantified or reproduced under exactly identical cellular conditions. Therefore, in-cell NMR analyses in bacteria, for example, are of qualitative nature. At low levels of recombinant protein over-expression, other cellular components become increasingly labeled, which results in the generation of signal artefacts that have to be actively suppressed (Rajagopalan et al., 2004). These adverse effects are absent when a labeled compound is introduced into the cellular environment at defined concentrations, as can be achieved by microinjection, for example. This approach, however, is restricted to a few eukaryotic cell types, which can be manipulated in such a way, and requires the labeled protein samples to be sufficiently soluble at high concentrations, since the maximum injection volume per single cell is typically on the order of nano-liters.

Alternatively, one could envisage intracellular sample delivery by means of cell-permeable synthetic vectors. We are particularly intrigued by the potential application of 'Trojan' peptide tags, which confer efficient cell membrane transduction activities to a wide range of fused protein substrates (Derossi et al., 1998; Dietz and Bahr, 2004). These internalization peptides are composed of short, positively charged amino acid residues, which can be genetically engineered to be part of virtually any recombinant polypeptide (Li et al., 2002). Upon labeled expression and purification of tagged fusion proteins, these substrates are simply added to the growth medium of a variety of cultured laboratory cell lines and readily internalized. In theory, this method should be generally applicable to a wide range of eukaryotic cells and quantitatively accomplishable for a large number of cells. The method of choice for eukaryotic in-cell NMR measurements will hence be dictated by both the suitability of the protein of interest for either approach and by the overall biological question to be addressed.

5.1. In-cell NMR spectroscopy in yeast-, insect- and mammalian cells

Recent advances in structural genomics have also led to a more thorough investigation of possible alternatives to bacterial recombinant protein expression and isotope labeling (Goto and Kay, 2000; Yokoyama, 2003). Amongst these, a few exotic approaches in mammalian CHO and HEK cells have been reported for NMR sample preparations (Hansen et al., 1992; Lustbader et al., 1996; Wyss et al., 1995). More prominent systems include the yeast *Pichia pastoris* and Baculovirus-infected insect cells (Chen et al., 2006; Pickford and O'Leary, 2004; Strauss et al., 2003, 2005). All of these eukaryotic cells have been employed to prepare labeled NMR samples for *in vitro* analyses and we can therefore only speculate about their experimental suitability for in-cell NMR measurements. The major obstacles for the selective labeling of recombinant proteins with NMR-active isotopes in eukaryotic cells have been the difficulty to achieve adequate levels of protein over-expression, sufficient isotope incorporation and the costs of isotope-enriched growth media. Growth media for NMR labeling in *E. coli* are simple in their composition, easily prepared and, depending on the type of labeling, relatively cheap. Bacteria will also incorporate isotopes with high efficiency (~98%). Labeling media for eukaryotic cells are sophisticated, they must often be obtained commercially for satisfactory results, and they are expensive. The yeast *P. pastoris* represents an exception to this notion since cells can be grown in glycerol/glucose and labeled in ^{15}N-ammoniumchloride and ^{13}C-methanol (Wood and Komives, 1999). Due to the complexity of most eukaryotic metabolisms, isotope incorporation in the above cells is typically less than 90%. With regard to in-cell NMR measurements, induction times for recombinant protein expression are on the order of days rather than hours, which is likely to increase the amount of background labeling artefacts.

In summary, we believe that in-cell NMR measurements in these eukaryotic cells will be technically feasible but are probably not going to constitute practicable routine approaches. Some proteins may require expression within a eukaryotic cellular environment in order to properly fold or to express in a soluble form. For these cases, in-cell NMR experiments in yeast, insect or mammalian cells may provide valuable insights into the structural mechanisms of protein folding or could enable the production of labeled, functional proteins for the delivery into other cells.

5.2. In-cell NMR spectroscopy in X. laevis oocytes

We, and others, have recently reported the usage of *X. laevis* oocytes for eukaryotic in-cell NMR measurements (Selenko et al., 2006; Serber et al., 2006; Sakai et al., 2006). These amphibian cells have long served as important laboratory tools in the disciplines of developmental and cellular biology (Fig. 4a) (Liu, 2006; Murray, 1991b). Mature oocyte cells (stage VI) arrest in prophase at the G2/M boundary of the first meiotic division (Fig. 4b) and contain large cell volumes (~1μL, compared to a few pL as for most somatic cells), 20% of which comprises the nuclear organelle (or germinal vesicle). During oocyte to egg maturation, a hormonal trigger activates synchronized cell cycle progression into metaphase of meiosis II. For isolated oocytes in an *in vitro* setting, these events can be executed by the external addition of hormones, which renders this system an important laboratory tool for studying signaling events during cell cycle progression. Cellular extracts from *Xenopus* oocytes or from *Xenopus* eggs are easily obtained in a virtually undiluted form and similarly recapitulate most of the intact cells' biological activities. They are frequently used as alternative, cell-free systems for *ex vivo* analyses of cellular processes (Crane and Ruderman, 2006; Murray, 1991a).

Stage VI oocytes are conveniently manipulated by microinjection, which permits the direct deposition of defined quantities of exogenous compounds into the cellular environment (Fig. 4c). Sophisticated setups and protocols for manipulating *Xenopus* oocytes have been devised over the years, which include a fully automated injection setup (Schnizler et al., 2003). We use this robotic device to routinely introduce labeled protein samples into large numbers of *Xenopus* oocytes for in-cell NMR measurements (Fig. 4d). About 200 manipulated oocytes are required for an in-cell NMR sample, which corresponds to >250 μL in sample volume and, inside a Shigemi™ NMR tube, will suffice to span transmitter- and receiver-coil extensions of most commercial NMR spectrometer probes (Fig. 4e). Considering the small injection volume per cell (50 nL), an oocyte sample requires only about 10 μL of labeled protein. The necessary concentration of injected protein for the minimally sufficient experimental

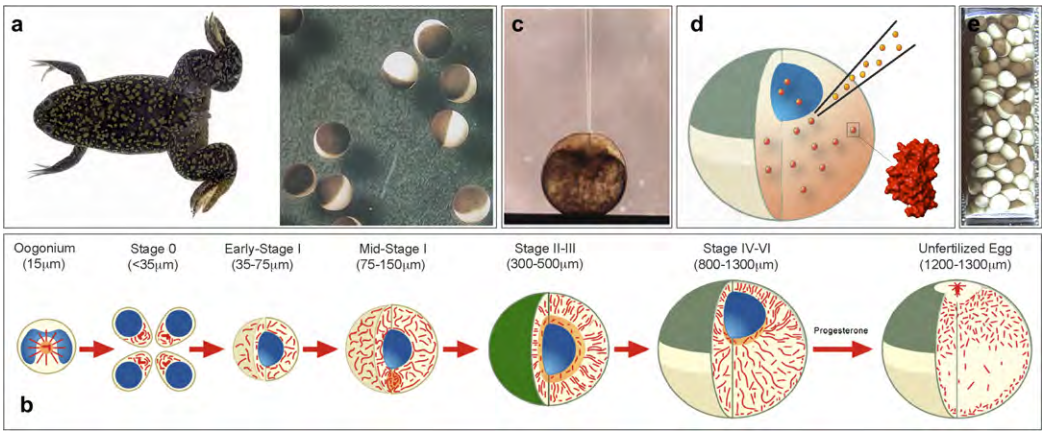

Fig. 4. (a) Oocyte cells from *Xenopus laevis* are surgically removed from the ovary lobes of female frogs. Stage VI oocytes are sorted and collected for microinjections. (b) An overview of the oocyte maturation process and oocyte to egg transition is depicted. The developmental stages and average cell sizes are indicated. Note that stage VI oocytes are interphase cells at the G2/M transition and contain an intact nuclear organelle (or germinal vesicle, shown in blue). Upon progesterone treatment, oocytes mature into eggs, which are arrested in metaphase of meiosis II. The rearrangement of cellular microtubules is shown in red. (Figure courtesy of David L. Gard, University of Utah). (c) Cell injections with pulled glass capillaries enable the quantitative deposition of exogenous proteins into the intracellular environment of *Xenopus* oocytes. (d) For in-cell NMR applications in *Xenopus* oocytes, the labeled protein sample is introduced into the unlabeled intracellular environment of intact oocyte cells by this approach. (e) The resulting in-cell NMR sample consists of ~200 oocytes settled in a Shigemi™ NMR tube.

NMR signal has to be evaluated empirically for each sample. In addition, the protein's size, expected cellular activity, and type of labeling, will determine the minimally sufficient amount of labeled compound. Injection concentrations are typically in the range of 0.5–3.0 mM. These quantities might not be achievable for all recombinant protein samples nor generally suffice for a satisfying experimental readout, especially for biomolecules with promiscuous binding activities towards endogenous cellular components, and within a reasonable amount of experimental time. It is evident that a compromise between the experimentally achievable signal-to-noise, the duration of individual NMR experiments, the cellular concentration of labeled proteins and its physiological relevance, will have to be found if in-cell NMR measurements in *X. laevis* oocytes are to yield biologically meaningful results.

We also conduct NMR experiments in *X. laevis* oocyte and egg extracts, and find that in these 'cellular' settings the overall spectral quality is generally better due to the homogenous nature of the sample solutions. Concentrations of labeled specimens in the lower micro-molar range (10–50 μM) typically yield interpretable 2-dimensional correlation spectra of satisfying quality. Additionally, sample dilutions upon extract suspension are lower than for oocyte injections. Most biological reactions are executed with comparable activities in *Xenopus* cell extracts and can easily be modulated by the addition of small molecule inhibitors or activators. These extracts can also be selectively depleted of certain cellular components, and replenished with labeled, NMR-active substitutes, which enables the isolated investigation of biological processes without compromising the qualitative nature of the experimental readout by residual endogenous activities. A more thorough description of experimental procedures is given in (Selenko et al., 2006; Serber et al., 2006)

Fig. 5a shows the characteristic appearance of 1-dimensional proton-only NMR spectra (no isotope-edited correlation 'filter') of a ^{15}N-labeled protein sample in its pure form (top panel), resuspended in crude *Xenopus* egg extract (middle panel) or upon oocyte injection (bottom panel). It is apparent that by conventionally recording all NMR signals from ^{1}H nuclei only, these spectra do not allow to discriminate between resonances from endogenous, cellular components and the ones from the labeled compound. Moreover, it is evident that the quality of in-cell NMR experiments, recorded in this mode, is too poor to provide any information on the injected protein sample. When, however, these same samples are measured with the application of a ^{1}H, ^{15}N correlation pulse-sequence, similar 1-dimensional traces selectively display NMR resonances from the labeled protein only. The good quality of these background-suppressed spectra enables the selective detection of protein NMR signals under extract and in-cell conditions (Fig. 5b). Changes in NMR line widths of individual resonance signals in these different aqueous solutions are readily visible (Fig. 3). Clearly, these spectra demonstrate the great potential of isotope-edited correlation techniques and the experimental feasibility of high-resolution NMR experiments in *Xenopus* egg extracts and oocyte cells.

We have recently delineated the experimental reference conditions for in-cell NMR measurements in *X. laevis* oocytes and provided a detailed structural analysis of the conformational *in vivo* properties of a small, biologically inert protein domain (Selenko et al., 2006). Whereas our sample did not engage in biological interactions in these cells, a similar approach by Sakai et al., presented experimental evidence for the cellular *in vivo* characteristics of ubiquitin, another small, but highly abundant protein with a large array of cellular functions (Sakai et al., 2006).

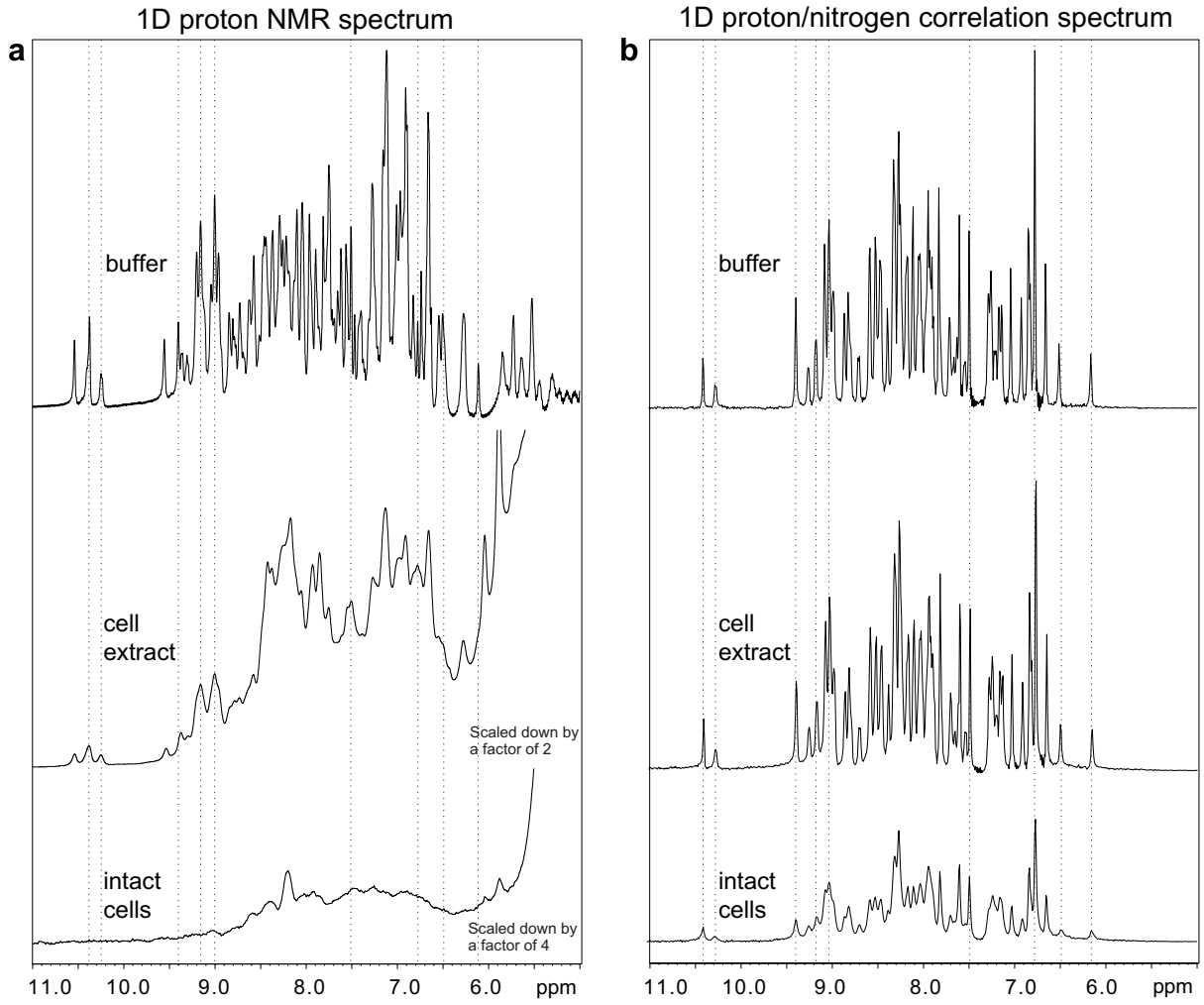

Fig. 5. (a) The amide/aromatic region (5–11 ppm) of a 1D proton spectrum of the *Streptococcal* protein GB1 (GB1) domain recorded with a standard ^{1}H NMR experiment. Solvent suppression was achieved by a Watergate NMR pulse-sequence, the experiment was recorded at 295 K, with 1024 points and 128 scans. The sample concentration was 0.5 mM of ^{15}N-labeled GB1 (No ^{15}N decoupling was employed). Individual panels display the respective NMR spectra in buffer (top), in crude *Xenopus* egg extracts (middle) and upon oocyte injection (bottom). Scaling was performed as indicated. (b) Identical samples as in (a) but recorded as 1D $^{1}H(^{15}N)$ hetero-nuclear single quantum coherence (HSQC) correlation spectra, using a Watergate version of a standard, sensitivity-enhanced HSQC pulse-sequence with the same number of scans as in (a). No scaling was applied to the experimental data. Some corresponding resonance frequencies are indicated as lines at the respective chemical shift values.

Indeed, given the multitude of cellular binding partners for ubiquitin in *Xenopus* oocytes, in-cell NMR experiments for the wild-type protein yielded poor quality NMR spectra. Only upon mutating conserved binding sites of known interaction surfaces, did ubiquitin provide interpretable in-cell NMR data. This work thus serves a fine example for what kind of biological studies will be amenable to in-cell NMR approaches. Once cellular proteins engage the labeled specimen in too many generic interactions, or scavenge the NMR-active protein into complexes of molecular weights too large for detection by conventional solution-state NMR methods, the envisaged in-cell NMR approach is likely to fail. In most instances, such unfavorable binding behaviors are easily pre-assessed by NMR experiments in cellular extracts (Serber et al., 2006). Nonetheless, and despite the fact that these considerations might be regarded as discouraging, even poor quality in-cell NMR experiments provide a wealth of information about the *in vivo* biological activity of a given protein and can still serve as valuable functional assays. In-cell NMR measurements in combination with mutant screening approaches, for example, can identify protein residues critically required for *in vivo* binding even if the nature of the biological activity that causes NMR signal deterioration is unknown. If, on the other hand, an expected cellular activity is highly selective, or engages the labeled protein of interest in a unique form of interaction, in-cell NMR experiments will succeed in providing a wealth of structural and functional insights. In the following chapter, we outline two biological areas of eukaryotic in-cell NMR research that we actively and successfully pursue in our laboratory.

6. Future applications

6.1. In-cell NMR analyses of intrinsically disordered proteins

Intrinsically disordered proteins (IDPs) represent a growing class of gene products (Dyson and Wright, 2005), which are characterized by lack of secondary and/or tertiary structure in their pure forms and at physiological pH (Uversky, 2002). IDPs are estimated to account for ~20% of all human proteins (Dunker et al., 2000) and exert important functions in key cellular processes (Dunker et al., 2002). A significant number of IDPs are implicated in human protein deposition diseases, in which a normally soluble polypeptide forms insoluble aggregates in a subset of cells and precipitates in the form of amyolid fibrils (Fink, 2005). Little is known about the general 3-dimensional conformation of IDPs in cellular environments and it is tempting to speculate whether unfolded protein conformations are preserved under native, intracellular conditions. Could it be possible that some IDPs are not *per se* structurally deficient protein entities but that their unfolded state results as a consequence of the isolated *in vitro* experimental setup employed for their characterization? In-cell NMR experiments on the natively unfolded FlgM protein indeed suggest a more folded conformation in the cellular environment of live bacteria (Dedmon et al., 2002). Are there cell type specific differences in the conformation of unfolded proteins and to what extent do different cellular environments modulate the pathological conversion of IDPs during amyloid formation? In-cell NMR spectroscopy appears to be a most appropriate tool to address these questions in live cells.

6.2. In situ observation of post-translational protein modifications

A limitation of in-cell NMR spectroscopy in bacteria is the inability to study post-translational protein modifications. Whereas the function of most eukaryotic proteins is regulated by a variety of sometimes transient, covalent modifications, mammalian proteins expressed in bacterial organisms are typically not modified. Eukaryotic protein modifications, with the exception of post-translational glycosylation, involve the covalent attachment of small chemical entities, acetyl-, methyl-, or phosphate-groups, which do not greatly alter the overall molecular weight of the modified substrates. This is particularly favorable for biomolecular NMR analyses as the spectral quality of a unmodified protein is likely to be preserved upon covalent modification. In addition, post-translational protein modifications greatly change the chemical environment of targeted residues, which translates into large chemical shift changes. In theory, transferring an unmodified protein substrate into a eukaryotic cellular environment should result in its post-translational modification by endogenous enzymes and according to a specific, biologically relevant pattern. In-cell NMR measurements should then enable the *in situ* observation of the establishment of these modifications in a time-dependent and residue-specific manner. We have recently confirmed that these *in vivo* approaches can indeed resolve cellular phosphorylation reactions in *X. laevis* oocytes (Selenko et al., submitted). They point to a plethora of possible in-cell NMR applications in eukaryotic post-translational protein modification research.

Acknowledgments

We thank Matthew Dedmon for carefully reading the manuscript. P.S. is a Human Science Frontier Program Organization (HSFPO) long-term fellow (LT00686/2004-C). This work was additionally supported by a National Institute of Health Grant GM47467 to G.W.

References

Bryant, J.E., 2006. In-cell protein dynamics. Mol. Biosyst. 2, 406–410.
Bryant, J.E., Lecomte, J.T., Lee, A.L., Young, G.B., Pielak, G.J., 2005. Protein dynamics in living cells. Biochemistry 44, 9275–9279.
Burz, D.S., Dutta, K., Cowburn, D., Shekhtman, A., 2006. Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). Nat. Methods 3, 91–93.
Chen, C.Y., Cheng, C.H., Chen, Y.C., Lee, J.C., Chou, S.H., Huang, W., Chuang, W.J., 2006. Preparation of amino-acid-type selective isotope labeling of protein expressed in *Pichia pastoris*. Proteins 62, 279–287.
Cohen, J.S., Lyon, R.C., Daly, P.F., 1989. Monitoring intracellular metabolism by nuclear magnetic resonance. Methods Enzymol. 177, 435–452.
Crane, R., Ruderman, J., 2006. Using *Xenopus* oocyte extracts to study signal transduction. Methods Mol. Biol., 435–445.
Dedmon, M.M., Patel, C.N., Young, G.B., Pielak, G.J., 2002. FlgM gains structure in living cells. Proc. Natl. Acad. Sci. USA 99, 12681–12684.
Derossi, D., Chassaing, G., Prochiantz, A., 1998. Trojan peptides: the penetratin system for intracellular delivery. Trends Cell Biol. 8, 84–87.
Dietz, G.P., Bahr, M., 2004. Delivery of bioactive molecules into the cell: the Trojan horse approach. Mol. Cell Neurosci. 27, 85–131.
Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., Obradovic, Z., 2002. Intrinsic disorder and protein function. Biochemistry 41, 6573–6582.
Dunker, A.K., Obradovic, Z., Romero, P., Garner, E.C., Brown, C.J., 2000. Intrinsic protein disorder in complete genomes. Genome. Inform. Ser. Workshop Genome. Inform. 11, 161–171.
Dyson, H.J., Wright, P.E., 2005. Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. 6, 197–208.
Etezady-Esfarjani, T., Herrmann, T., Horst, R., Wuthrich, K., 2006. Automated protein NMR structure determination in crude cell-extract. J. Biomol. NMR 34, 3–11.
Fink, A.L., 2005. Natively unfolded proteins. Curr. Opin. Struct. Biol. 15, 35–41.
Fischer, M.W., Zeng, L., Majumdar, A., Zuiderweg, E.R., 1998. Characterizing semilocal motions in proteins by NMR relaxation studies. Proc. Natl. Acad Sci. USA 95, 8016–8019.
Goto, N.K., Kay, L.E., 2000. New developments in isotope labeling strategies for protein solution NMR spectroscopy. Curr. Opin. Struct. Biol. 10, 585–592.
Hansen, A.P., Petros, A.M., Mazar, A.P., Pederson, T.M., Rueter, A., Fesik, S.W., 1992. A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells. Biochemistry 31, 12713–12718.
Hubbard, J.A., MacLachlan, L.K., King, G.W., Jones, J.J., Fosberry, A.P., 2003. Nuclear magnetic resonance spectroscopy reveals the

functional state of the signalling protein CheY *in vivo* in *Escherichia coli*. Mol. Microbiol. 49, 1191–1200.

Li, Y., Rosal, R.V., Brandt-Rauf, P.W., Fine, R.L., 2002. Correlation between hydrophobic properties and efficiency of carrier-mediated membrane transduction and apoptosis of a p53 C-terminal peptide. Biochem. Biophys. Res. Commun. 298, 439–449.

Liu, J.X., 2006. *Xenopus* Protocols, vol. 322.

Lustbader, J.W., Birken, S., Pollak, S., Pound, A., Chait, B.T., Mirza, U.A., Ramnarain, S., Canfield, R.E., Brown, J.M., 1996. Expression of human chorionic gonadotropin uniformly labeled with NMR isotopes in Chinese hamster ovary cells: an advance toward rapid determination of glycoprotein structures. J. Biomol. NMR 7, 295–304.

McNulty, B.C., Young, G.B., Pielak, G.J., 2006. Macromolecular crowding in the *Escherichia coli* periplasm maintains alpha-synuclein disorder. J. Mol. Biol. 355, 893–897.

Murray, A.W., 1991a. Cell cycle extracts. Methods Cell Biol. 36, 581–605.

Murray, A.W., 1991b. *Xenopus laevis*: practical uses in cell and molecular biology. Methods Cell Biol. 36, 1–718.

Palmer 3rd, A.G., 2001. NMR probes of molecular dynamics: overview and comparison with other techniques. Annu. Rev. Biophys. Biomol. Struct. 30, 129–155.

Peng, J.W., Wagner, G., 1994. Investigation of protein motions via relaxation measurements. Methods Enzymol. 239, 563–596.

Pickford, A.R., O'Leary, J.M., 2004. Isotopic labeling of recombinant proteins from the methylotrophic yeast *Pichia pastoris*. Methods Mol. Biol. 278, 17–33.

Rajagopalan, S., Chow, C., Raghunathan, V., Fry, C.G., Cavagnero, S., 2004. NMR spectroscopic filtration of polypeptides and proteins in complex mixtures. J. Biomol. NMR 29, 505–516.

Reardon, P.N., Spicer, L.D., 2005. Multidimensional NMR spectroscopy for protein characterization and assignment inside cells. J. Am. Chem. Soc. 127, 10848–10849.

Reckel, S., Lohr, F., Dotsch, V., 2005. In-cell NMR spectroscopy. Chembiochem 6, 1601–1606.

Sakai, T., Tochio, H., Tenno, T., Ito, Y., Kokubo, T., Hiroaki, H., Shirakawa, M., 2006. In-cell NMR spectroscopy of proteins inside *Xenopus laevis* oocytes. J. Biomol. NMR 36, 179–188.

Schnizler, K., Kuster, M., Methfessel, C., Fejtl, M., 2003. The roboocyte: automated cDNA/mRNA injection and subsequent TEVC recording on *Xenopus oocytes* in 96-well microtiter plates. Receptors Channels 9, 41–48.

Selenko, P., Serber, Z., Gadea, B., Ruderman, J., Wagner, G., 2006. Quantitative NMR analysis of the protein GB1 domain in *Xenopus laevis* egg extracts and intact oocytes. Proc. Natl. Acad. Sci. USA 32, 11904–11909.

Serber, Z., Corsini, L., Durst, F., Dotsch, V., 2005. In-cell NMR spectroscopy. Methods Enzymol. 394, 17–41.

Serber, Z., Keatinge-Clay, A.T., Ledwidge, R., Kelly, A.E., Miller, S.M., Dotsch, V., 2001. High-resolution macromolecular NMR spectroscopy inside living cells. J. Am. Chem. Soc. 123, 2446–2447.

Serber, Z., Selenko, P., Hansel, R., Reckel, S., Lohr, F., Ferrell, J.E., Wagner, G., Dotsch, V., 2006. Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy. Nat. Protocols 1, 2701–2709.

Serber, Z., Straub, W., Corsini, L., Nomura, A.M., Shimba, N., Craik, C.S., Ortiz de Montellano, P., Dotsch, V., 2004. Methyl groups as probes for proteins and complexes in in-cell NMR experiments. J. Am. Chem. Soc. 126, 7119–7125.

Strauss, A., Bitsch, F., Cutting, B., Fendrich, G., Graff, P., Liebetanz, J., Zurini, M., Jahnke, W., 2003. Amino-acid-type selective isotope labeling of proteins expressed in Baculovirus-infected insect cells useful for NMR studies. J. Biomol. NMR 26, 367–372.

Strauss, A., Bitsch, F., Fendrich, G., Graff, P., Knecht, R., Meyhack, B., Jahnke, W., 2005. Efficient uniform isotope labeling of Abl kinase expressed in Baculovirus-infected insect cells. J. Biomol. NMR 31, 343–349.

Szwergold, B.S., 1992. NMR spectroscopy of cells. Annu. Rev. Physiol. 54, 775–798.

Uversky, V.N., 2002. What does it mean to be natively unfolded? Eur. J. Biochem. 269, 2–12.

Wieruszeski, J.M., Bohin, A., Bohin, J.P., Lippens, G., 2001. In vivo detection of the cyclic osmoregulated periplasmic glucan of *Ralstonia solanacearum* by high-resolution magic angle spinning NMR. J. Magn. Reson. 151, 118–123.

Wood, M.J., Komives, E.A., 1999. Production of large quantities of isotopically labeled protein in *Pichia pastoris* by fermentation. J. Biomol. NMR 13, 149–159.

Wyss, D.F., Choi, J.S., Li, J., Knoppers, M.H., Willis, K.J., Arulanandam, A.R., Smolyar, A., Reinherz, E.L., Wagner, G., 1995. Conformation and function of the N-linked glycan in the adhesion domain of human CD2. Science 269, 1273–1278.

Yokoyama, S., 2003. Protein expression systems for structural genomics and proteomics. Curr. Opin. Chem. Biol. 7, 39–43.

Zuiderweg, E.R., 2002. Mapping protein–protein interactions in solution by NMR spectroscopy. Biochemistry 41, 1–7.