

# Isotope labeling methods for studies of excited protein states by relaxation dispersion NMR spectroscopy

Patrik Lundström<sup>1</sup>, Pramodh Vallurupalli<sup>2</sup>, D Flemming Hansen<sup>2</sup> & Lewis E Kay<sup>2</sup>

<sup>1</sup>Division of Molecular Biotechnology, Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden. <sup>2</sup>Departments of Molecular Genetics, Biochemistry and Chemistry, The University of Toronto, Toronto, Ontario, Canada. Correspondence should be addressed to L.E.K. (kay@pound.med.utoronto.ca).

Published online 22 October 2009; doi:10.1038/nprot.2009.118

**The utility of nuclear magnetic resonance (NMR) spectroscopy as a tool for the study of biomolecular structure and dynamics has benefited from the development of facile labeling methods that incorporate NMR active probes at key positions in the molecule. Here we describe a protocol for the labeling of proteins that facilitates their study using a technique that is sensitive to millisecond conformational exchange processes. The samples necessary for an analysis of exchange dynamics are discussed, using the Abp1p SH3 domain from *Saccharomyces cerevisiae* as an example. For this system, the time frame for production of each sample, including *in vitro* refolding, is about 80 h. The samples so produced facilitate the measurement of accurate chemical shifts of low populated, invisible conformers that are part of the exchange pathway. The accuracy of the methodology has been established experimentally and the chemical shifts that are obtained provide important restraints in structure calculations of the excited state.**

## INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool for the study of biomolecular dynamics spanning a wide range of time scales<sup>1,2</sup>. Central to the utility of NMR technique has been the ability to label the sample of interest with the appropriate NMR probes (such as <sup>15</sup>N, <sup>13</sup>C and <sup>2</sup>H or in some cases <sup>1</sup>H in a deuterated background)<sup>3</sup> and the subsequent design of sophisticated multi-dimensional experiments<sup>4</sup>, which exploit the labeling patterns chosen. Building on advances in both labeling methodologies, spectrometer performance and NMR experiments, the Carr–Purcell–Meiboom–Gill (CPMG)-relaxation dispersion NMR experiment, which dates back to ideas first formulated a half century ago<sup>5,6</sup>, has emerged as a powerful method for characterizing protein dynamics on the millisecond time scale. This regime is of particular relevance, as it encompasses a wide range of biological processes including protein folding<sup>7–10</sup>, enzyme catalysis<sup>11–16</sup> and ligand binding<sup>10,17</sup>. In the case of a protein undergoing a millisecond time-scale exchange process described by  $A \xrightleftharpoons[k_{BA}]{k_{AB}} B$ ,

where *B* is a low populated, often invisible state, the information content of such experiments is the exchange rate,  $k_{ex} = k_{AB} + k_{BA}$ , the population of the minor state,  $p_B$  and the absolute value of the chemical shift difference between states *A* and *B*,  $|\Delta\omega|$  (ref. 18). The chemical shift of the invisible state can be obtained from  $|\Delta\omega|$ , as the sign of  $\Delta\omega$  is available from a set of additional experiments that have been described previously<sup>19</sup>. Thus, this methodology allows one to ‘see the invisible’ in the sense that information about state *B* can be obtained, which in general is lacking in studies using other biophysical techniques.

To this point, a considerable focus of many applications has been on the characterization of the thermodynamic and kinetic properties of the exchange process through the analysis of parameters  $k_{ex}$  and  $p_B$  (ref. 20). However, it is clear that valuable information is also available from chemical shifts. For example, it has long been recognized that the chemical shift is sensitive to protein conformation, with different nuclei reporting on different aspects of structure<sup>21–23</sup>. Knowledge of chemical shifts of the excited conformer through

CPMG experiments thus opens up an avenue for the detailed structural characterization of this elusive state. Particularly interesting in this regard is the recent demonstration that the chemical shifts of six different nuclei, <sup>13</sup>C $\alpha$ , <sup>13</sup>C $\beta$ , <sup>13</sup>CO, <sup>1</sup>H $\alpha$ , <sup>1</sup>H<sup>N</sup> and <sup>15</sup>N, can be used as constraints in structural calculations of the ground states of small proteins<sup>24,25</sup> to produce structures of a quality that is comparable with those calculated using the more traditional NMR constraints, including nuclear Overhauser enhancements (NOEs) and residual dipolar couplings (RDCs)<sup>26–30</sup>. The chemical shifts listed above will be equally useful in structural studies of protein excited states assuming that they can be determined accurately by CPMG dispersion experiments<sup>31</sup>. In this regard the choice of labeling is critical. Uniform <sup>13</sup>C or <sup>1</sup>H labeling can lead to significant artifacts in <sup>13</sup>C or <sup>1</sup>H CPMG data sets, respectively, and potentially to substantial errors in <sup>13</sup>C or <sup>1</sup>H chemical shifts of the excited state. The major sources of artifacts are homonuclear scalar couplings, which would, of course, be removed or at the very least minimized through the development of suitable selective labeling protocols. Robust labeling methods (and the associated NMR experiments) are now in place for measuring chemical shifts of protein excited states using relaxation dispersion methods with an accuracy level that far exceeds that at which these shifts can be predicted<sup>32</sup> (see below), paving the way for their use in structural calculations. The purpose of this paper is to describe protocols for the preparation of suitable NMR samples that are needed for the measurement of chemical shifts by relaxation dispersion measurements. We note in passing that these samples are also of use in recording additional excited-state structural restraints based on the orientation of bond vectors with respect to a molecule fixed frame<sup>33,34</sup>, and that they can also be used in standard heteroatom relaxation experiments (*T*<sub>1</sub>, *T*<sub>2</sub> and heteronuclear NOE) for characterization of ps–ns time-scale dynamics of ground states of proteins<sup>35</sup> at <sup>15</sup>N or isolated <sup>13</sup>C $\alpha$  and <sup>13</sup>C $\beta$  positions.

## Overview of the approach

A labeling scheme should optimally meet as many of the following criteria as possible: (i) it should lead to high levels of enrichment



**TABLE 1** | Summary of protein samples required for relaxation dispersion experiments.

Sample	Observed nucleus <sup>a</sup>	Carbon source(s)	Growth solvent	Bacterial strain
1	<sup>15</sup> N, <sup>1</sup> H <sup>N</sup> and <sup>13</sup> CO	[ <sup>13</sup> C <sub>6</sub> , <sup>2</sup> H <sub>7</sub> ]-glucose	100% D <sub>2</sub> O <sup>b</sup>	BL21(DE3)
2	<sup>1</sup> H $\alpha$ (18)	[ <sup>13</sup> C <sub>6</sub> , <sup>2</sup> H <sub>7</sub> ]-glucose	50% D <sub>2</sub> O/50% H <sub>2</sub> O <sup>c</sup>	BL21(DE3)
3	<sup>13</sup> C $\alpha$ (17)	[2- <sup>13</sup> C]-glucose	100% H <sub>2</sub> O <sup>d</sup>	BL21(DE3)
4	<sup>13</sup> C $\beta$ (11)	[1- <sup>13</sup> C]-glucose/NaHCO <sub>3</sub> (nat. ab.)	100% H <sub>2</sub> O <sup>d</sup>	BL21(DE3) $\Delta$ sdh <sup>e</sup>
5	<sup>13</sup> C $\beta$ (4)	[2- <sup>13</sup> C]-glucose	100% H <sub>2</sub> O <sup>d</sup>	BL21(DE3) $\Delta$ sdh <sup>e</sup>

<sup>a</sup>Numbers of residue-types for which chemical shifts can be obtained using existing NMR methodology are indicated in parentheses. <sup>b</sup>The NMR experiments must be carried out in H<sub>2</sub>O with all deuterons at the amide positions exchanged to protons. <sup>c</sup>The NMR experiments must be carried out in D<sub>2</sub>O with all protons at the amide positions exchanged to deuterons. <sup>d</sup>The NMR experiments can be carried out in either H<sub>2</sub>O or D<sub>2</sub>O. <sup>e</sup>Other bacterial strains with a disrupted TCA cycle, like DL323, can also be used although this may require recloning of the target gene<sup>36</sup>. The BL21(DE3) $\Delta$ sdh strain is described in ref. 38.

at only the desired positions; (ii) it should work for all residue-types; (iii) it should be compatible with high levels of protein overexpression; (iv) it should be simple to implement; (v) it should minimize the number of different samples that need to be prepared; (vi) it should be cost effective in terms of the isotopically labeled compounds used. Unfortunately, many of these criteria are difficult if not impossible to simultaneously achieve and the labeling strategy must often necessarily be a compromise between the various criteria listed above. To prepare the samples needed for relaxation dispersion experiments measuring <sup>13</sup>C $\alpha$ , <sup>13</sup>C $\beta$ , <sup>13</sup>CO, <sup>1</sup>H $\alpha$ , <sup>1</sup>H<sup>N</sup> and <sup>15</sup>N chemical shifts, we use an approach whereby suitably labeled glucose is the main carbon source. This maximizes bacterial growth and yield of protein while keeping the expression protocols simple. Glucose that is suitably isotopically enriched for a large number of different applications is also commercially available, often at reasonable cost. The main disadvantage of glucose compared with other general carbon sources, such as selectively labeled glycerol<sup>36</sup> or pyruvate<sup>37</sup>, is lower levels of label incorporation (see ANTICIPATED RESULTS).

We propose the five samples presented in **Table 1** for recording CPMG dispersion profiles from which the chemical shifts listed above are extracted. In all cases overexpression is carried out in the *Escherichia coli* strain BL21(DE3), or a derivative of this strain with a dysfunctional TCA cycle, BL21(DE3) $\Delta$ sdh<sup>38</sup>, using M9 minimal medium, which is supplemented with <sup>15</sup>NH<sub>4</sub>Cl as the nitrogen source and the appropriately labeled glucose as the carbon source. The growth solvent is either 100% H<sub>2</sub>O, 100% D<sub>2</sub>O or a mixture of the two.

**Sample 1:** uniformly <sup>13</sup>C/<sup>15</sup>N/<sup>2</sup>H labeled protein with amide positions back exchanged to protons. This sample can be used for <sup>15</sup>N (refs. 39,40), <sup>1</sup>H<sup>N</sup> (ref. 41) and <sup>13</sup>CO (refs. 42,43)-relaxation dispersion experiments. Chemical shifts in the excited state are obtained for all residue-types except for Pro (<sup>15</sup>N shifts) and residues preceding Pro (<sup>13</sup>CO shifts), as Pro does not have an attached amide proton that is required in the experiments.

**Sample 2:** uniformly <sup>15</sup>N/<sup>13</sup>C labeled and fractionally deuterated protein. This sample is used for <sup>1</sup>H $\alpha$  dispersion experiments<sup>44</sup>. By carrying out overexpression using deuterated glucose as the carbon source in 50% D<sub>2</sub>O, the H $\alpha$  positions will be 50% deuterated, whereas all other positions on average will have higher deuteration levels<sup>44</sup>. Here it is essential that the amide positions are deuterated,

which is achieved by carrying out the experiments in D<sub>2</sub>O. All residue-types are appropriately labeled, but for technical reasons Ser and Thr can not in general be analyzed using the proposed experiments<sup>44</sup>. It should be noted that we prefer to use this sample instead of Sample 3 for the measurement of <sup>13</sup>C $\alpha$ -relaxation dispersion profiles of Gly<sup>45</sup>.

**Sample 3:** uniformly <sup>15</sup>N, fractionally <sup>13</sup>C labeled protein. By carrying out overexpression in [2-<sup>13</sup>C]-glucose most residue-types will be enriched for <sup>13</sup>C at C $\alpha$  but not at C $\beta$  and CO positions<sup>46</sup>. For solvent suppression purposes the CPMG experiments can be recorded in D<sub>2</sub>O, but otherwise this is not essential for carrying out the experiment. Accurate <sup>13</sup>C $\alpha$  dispersion experiments, without the need for purge elements to filter out contributions due to <sup>13</sup>C-<sup>13</sup>C spin pairs<sup>31</sup>, can be recorded for 17 residue-types. Leu which is not significantly enriched at C $\alpha$ , and Ile and Val that are simultaneously <sup>13</sup>C $\alpha$ /<sup>13</sup>C $\beta$  labeled are the exceptions.

**Sample 4:** uniformly <sup>15</sup>N and fractionally <sup>13</sup>C labeled protein. The C $\beta$ , but not the C $\alpha$  or C $\gamma$ , positions are enriched in <sup>13</sup>C for 11 different residue-types by using [1-<sup>13</sup>C]-glucose as the carbon source during overexpression<sup>38</sup>. The generation of isolated <sup>13</sup>C $\beta$  spins is maximized through the use of BL21(DE3) $\Delta$ sdh cells that are deficient in succinate dehydrogenase, thereby avoiding scrambling in the TCA cycle, and by the addition of NaHCO<sub>3</sub> (natural abundance) that avoids formation of <sup>13</sup>C-<sup>13</sup>C bonds when phosphoenolpyruvate is carboxylated (that ultimately will produce <sup>13</sup>C $\beta$ -<sup>13</sup>C $\gamma$  carbon pairs for Asn, Asp, Lys, Met and Thr)<sup>38</sup>.

**Sample 5:** uniformly <sup>15</sup>N and fractionally <sup>13</sup>C labeled protein. C $\beta$ , but not C $\alpha$  or C $\gamma$ , positions are enriched in <sup>13</sup>C for four different residue-types by using [2-<sup>13</sup>C]-glucose as the carbon source during overexpression<sup>38</sup>. BL21(DE3) $\Delta$ sdh cells that have a disrupted TCA cycle are used to avoid scrambling and to maximize yields.

Although <sup>15</sup>N labeling is not necessary for samples used for recording <sup>1</sup>H $\alpha$ , <sup>13</sup>C $\alpha$ - and <sup>13</sup>C $\beta$ -relaxation dispersion experiments, it is convenient to include this isotope, as it does not interfere with the experiments being carried out and provides an inexpensive way of checking the integrity of the protein by <sup>15</sup>N-<sup>1</sup>H<sup>N</sup> correlation spectroscopy. As all dispersion experiments on Sample 1 involve amide proton detection, it is of course necessary that all deuterons at the amide positions are exchanged to protons (as protein

production is in D<sub>2</sub>O, all amides will initially be deuterated) and that the NMR experiments be carried out in H<sub>2</sub>O. To obtain a complete exchange at all positions it is often necessary to unfold the protein with subsequent refolding in H<sub>2</sub>O. Unfolding is typically carried out in 6-M GdnHCl, which is removed by rapid dilution in and/or dialysis against buffers that favor the native conformation<sup>47</sup>. To minimize aggregation it is often necessary to add stabilizing agents such as arginine<sup>48</sup>, and for proteins containing disulfide bridges addition of a redox couple is often necessary<sup>49</sup>. Similarly, the experi-

ments involving all other samples are optimally carried out in D<sub>2</sub>O, and for Sample 2 it is important that deuterons be substituted for protons at the amide positions. Unfortunately, unfolding and refolding in D<sub>2</sub>O would be very expensive because of the large volumes involved. We thus prefer to carry out all protein purification steps in H<sub>2</sub>O and then dissolve lyophilized protein in a small volume of D<sub>2</sub>O (with heating or elevated pH if possible to accelerate the H-D exchange). Note that in some cases not all positions will be fully exchanged.

## MATERIALS

### REAGENTS

- Isopropyl-beta-D-1-thiogalactopyranoside (IPTG)
- LB rich growth medium
- M9 minimal medium (see below)
- BL21(DE3) cells
- BL21(DE3)Δsdh cells (available from the authors upon request)
- Plasmid containing gene-of-interest under control of the T7 promoter
- Isotopically labeled glucose
- D<sub>2</sub>O **! CAUTION** Do not ingest.
- Tobacco etch virus (TEV) protease (with isolated His<sub>6</sub> tag)
- Lysis/denaturing buffer (6-M GdnHCl, 0.1-M sodium phosphate and 2-mM imidazole, pH 8.0) **! CAUTION** Irritates skin and eyes; wear protective gloves.
- Elution buffer (6-M GdnHCl and 0.2-M acetic acid) **! CAUTION** Irritates skin and eyes; wear protective gloves.
- Refolding buffer 1 (10-mM TrisHCl, 1-M NaCl and 4-mM DTT, pH 8.0)
- Refolding buffer 2 (10-mM TrisHCl, 250-mM NaCl and 4-mM DTT, pH 8.0)
- Gel filtration buffer (50-mM sodium phosphate, 100-mM NaCl and 1-mM EDTA, pH 7.0)
- Na<sub>3</sub>

### EQUIPMENT

- Sterile flasks for cell growth
- Incubating shaker for bacterial growth
- Preparative centrifuge
- UV/Vis spectrophotometer
- Purification columns and protein concentrators

### REAGENT SETUP

**M9 minimal medium** For 1 liter of growth medium, mix together 6-g Na<sub>2</sub>HPO<sub>4</sub>, 3-g KH<sub>2</sub>PO<sub>4</sub> and 0.5-g NaCl, 0.5-g <sup>15</sup>NH<sub>4</sub>Cl, 11-mg CaCl<sub>2</sub> and 120-mg MgSO<sub>4</sub>. Add H<sub>2</sub>O (or D<sub>2</sub>O or a mixture of the two, see **Table 1**) to a final volume of 1 liter. Adjust pH to 7.4 and filter sterilize. Immediately before the start of the culture, take out 20 ml of the medium and to this add 3-g

glucose, 10-mg biotin, 10-mg thiamin and the appropriate antibiotic(s). For the production of Sample 4 include also NaHCO<sub>3</sub> (natural abundance) so that the final concentration in the 1 liter of culture will be 20 mM. Filter sterilize and add to the culture. In the case of Sample 4, set aside an additional stock solution of NaHCO<sub>3</sub> (filter sterilized), which will be added to the 1 liter culture in a subsequent step to increase the concentration of NaH<sup>12</sup>CO<sub>3</sub> by 10 mM (see below). The media should be used within 24 h of adding glucose, vitamins and antibiotics. Filter sterilized (or autoclaved) salts can be stored at room temperature (25 °C) for up to 6 months.

**Isotopically labeled compounds** As listed in **Table 1**, the sole nitrogen source for growth and overexpression for all samples is <sup>15</sup>NH<sub>4</sub>Cl, whereas both the carbon source and the solvent vary depending on the sample. For Samples 1 and 2, [<sup>13</sup>C<sub>6</sub>, <sup>2</sup>H<sub>7</sub>]-glucose is used as carbon source, [2-<sup>13</sup>C]-glucose for Samples 3 and 5, whereas Sample 4 requires [1-<sup>13</sup>C]-glucose. The solvent for culture growths is 100% D<sub>2</sub>O, 50% D<sub>2</sub>O/50% H<sub>2</sub>O and 100% H<sub>2</sub>O for Samples 1, 2 and 3–5, respectively.

**Cost** The cost per sample will depend on the protein yield per liter growth medium and on the particular labeling scheme. Based on the following list prices for isotopes and the recommended amounts of each isotope, we estimate the total cost of producing different samples per liter of growth medium to be: \$770 for Sample 1, \$615 for Sample 2, \$1,095 for Samples 3 and 5 and \$360 for Sample 4.

[1- <sup>13</sup> C] glucose	\$115 g <sup>-1</sup>
[2- <sup>13</sup> C]-glucose	\$360 g <sup>-1</sup>
[ <sup>13</sup> C <sub>6</sub> ]-glucose	\$68 g <sup>-1</sup>
[ <sup>13</sup> C <sub>6</sub> , <sup>2</sup> H <sub>7</sub> ]-glucose	\$150 g <sup>-1</sup>
<sup>15</sup> NH <sub>4</sub> Cl	\$25 g <sup>-1</sup>
D <sub>2</sub> O (99%)	\$305 liter <sup>-1</sup>

## PROCEDURE

### Protein overexpression in labeled minimal medium (M9) for the biosynthesis of any of the five isotopically enriched samples listed above

- 1| Transfer one or more freshly transformed *E. coli* colonies to 30-ml LB (in H<sub>2</sub>O) supplemented with the appropriate antibiotic(s) and grow cells at 37 °C in a shaking incubator until OD<sub>600</sub> = 1.0 is reached.
- 2| Spin down the cells at 1,200g, 15 min at room temperature.
- 3| Resuspend a fraction of cells in 10% of the isotopically labeled M9 medium to achieve OD<sub>600</sub> of 0.1–0.2. This starter culture, including growth solvent, has the identical isotopic composition as the main culture. Grow the cells at 37 °C until OD<sub>600</sub> = 1.0. Pour the starter culture directly into the remaining 90% of the isotopically enriched M9 medium (900 ml for a 1 liter culture).
- 4| Grow the cells at 37 °C until OD<sub>600</sub> = 0.6–1.0.
- 5| Induce overexpression with 1-mM IPTG. For Sample 4 also add an additional 10-mM NaHCO<sub>3</sub> (natural abundance). Perform overexpression either at 37 °C for 2–5 h or at room temperature or 16 °C overnight. The final OD<sub>600</sub> will depend on the growth medium and on the duration of overexpression.

## PROTOCOL

▲ **CRITICAL STEP** The optimal temperature and duration of overexpression are highly protein specific and depend on the stability of the protein. Furthermore, selective labeling strategies are also expected to work better if the overexpression period is kept as short as possible so as to minimize scrambling. For the samples considered here, overnight expression has proven to be satisfactory.

6| Harvest cells by centrifugation at 1,200*g*, for 15 min at 4 °C and purify according to protein specific protocol.

### **A purification protocol including *in vitro* refolding and proteolytic cleavage of a thioredoxin–His<sub>6</sub> fusion tag (optimized for the Abp1p SH3 domain from *Saccharomyces cerevisiae*)**

7| Lyse cells by sonication on ice for 4 min in 60 ml in lysis/denaturing buffer.

8| Spin down cell debris at 13,000*g* for 30 min at 4 °C and pass supernatant through a 0.45- $\mu$ m filter.

9| Bind the protein to a HisTrap column (GE Healthcare, Biosciences AB, Uppsala, Sweden) equilibrated with lysis/denaturing buffer and wash with the same buffer until no more protein comes off.

10| Elute the protein with elution buffer.

11| Measure  $A_{280}$  and dilute the protein to  $<0.5$  mg ml<sup>-1</sup> in elution buffer.

12| Refold the protein by dialysis against 4-liter refolding buffer 1 and then 4-liter refolding buffer 2.

▲ **CRITICAL STEP** The refolding protocol is highly protein specific and must be optimized for every protein; see text.

13| Spin down any precipitate and pass supernatant through a 0.45- $\mu$ m filter. Do this even if there is no visible precipitate. Continue with the purification protocol. To maximize yields, any precipitate can be dissolved in lysis/denaturing buffer and refolded once more.

14| Cleave off the thioredoxin–His<sub>6</sub> tag with TEV protease (1 mg liter<sup>-1</sup> culture) at room temperature for 6 h and check that cleavage is complete on a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel. If not, add more TEV protease and continue cleavage until completion.

15| Apply to HisTrap column the thioredoxin–His<sub>6</sub> and TEV protease (which also has a His<sub>6</sub> tag), which will bind to the column, whereas the SH3 domain will come out in the flow-through.

16| Concentrate the flow-through from the HisTrap column containing the SH3 domain in concentrators with a 3-kDa cut-off to a volume of 2 ml. Pass the protein through a 0.45- $\mu$ m filter.

17| Apply to a Superdex S75 (GE Healthcare) column equilibrated with gel filtration buffer.

18| Analyze the fractions on SDS-PAGE gel and pool all fractions containing pure SH3 domain. Concentrate to a final concentration of 1 mM in concentrators with a 3-kDa cut-off.

19| Prepare the NMR sample by adding NaN<sub>3</sub> to a concentration of 1 mM. For Sample 1, add 10% D<sub>2</sub>O/NMR buffer (volume fraction) and transfer 500  $\mu$ l to an NMR tube. Lyophilize the other samples and dissolve the protein in 500- $\mu$ l 99% D<sub>2</sub>O before transferring to NMR tubes.

### ● **TIMING**

In a typical application cells are transformed and plated in the evening. The next morning colonies are transferred to an LB liquid culture and grown for 3 h (Step 1). The rest of day is spent growing the starter culture followed by the main culture (Steps 2–4). Overexpression is induced in the evening (Step 5) and the bacteria are harvested next morning (Step 6). The aggregate time for these steps will thus typically be 40 h. Purification including refolding and preparation of the NMR sample (Steps 7–19) can typically be done on a similar time-scale although this depends on both the purification and refolding protocols. The minimal total time required for obtaining an NMR sample is thus 80 h.

### **ANTICIPATED RESULTS**

As we propose to use glucose as the main carbon source, our protocols are designed to maximize protein yields. However, the yield per unit volume of growth medium will depend critically on the nature of the protein and on the expression system.

Even for small proteins that are apparently very similar, the differences can be large. For instance, we routinely get five 1-mM NMR samples per liter of growth medium for the Abp1p SH3 domain ( $\approx 20$ -mg protein), whereas the Fyn SH3 domain yields only one such sample per liter growth medium despite the fact that both proteins are expressed in BL21(DE3) cells, protein is purified using identical protocols and the two proteins share 36% sequence identity. Protein yields do drop for samples expressed in BL21(DE3) $\Delta$ sdh cells, but only by 25% relative to the 'normal' BL21(DE3) cell line<sup>38</sup>. Another critical aspect that determines yields is the refolding step required for some of the samples listed in **Table 1**. In general the refolding efficiency is highly protein specific, and may be inefficient or impossible even for some small proteins, especially those containing multiple disulfide bridges. In addition to the total amount of protein produced, the level of incorporation of label at the desired sites is also an important factor. In some cases, such as when  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$  or  $^1\text{H}\alpha$  chemical shifts in the excited state are to be probed by relaxation dispersion experiments, this label must be 'isolated' so that the deleterious effects brought about by homonuclear scalar couplings are minimized. Sample 1 is produced with uniform  $^{13}\text{C}$  and  $^{15}\text{N}$  labeling and with minimal incorporation of protons at positions other than the amide sites. Thus,  $^{15}\text{N}$ ,  $^1\text{H}$  and  $^{13}\text{CO}$  probes that are exploited in experiments being carried out on this sample are at a level of occupancy of very close to 100% (although amide deuterons that are present because the growth solvent is  $\text{D}_2\text{O}$  must be replaced by protons). All probes are available for study in all 20 residue-types using the experiments that are currently in place, with the exception of  $^{15}\text{N}$  for Pro and  $^{13}\text{CO}$  for all residues that precede Pro. Note that the presence of  $^{13}\text{CO}$ - $^{13}\text{C}\alpha$  spin pairs does not interfere with the recording of  $^{13}\text{CO}$  chemical shifts, as it is possible to refocus  $^{13}\text{CO}$  spins selectively using appropriate band-selective pulses<sup>42,43</sup>.

In the case of Sample 2 all residue-types are expected to be 50% protonated at the  $\alpha$  position with the level of deuteration at  $\text{C}\beta$ , ranging from 50 to 90%, depending on the amino acid (**Table 2**)<sup>50</sup>. The comparison of signal intensities of different isotopomers in  $^{13}\text{C}$ - $^1\text{H}$  correlation spectra supports these predictions<sup>44</sup>. The loss of half of the protons at  $\alpha$  positions will reduce the signal-to-noise in experiments by 50%, but this is necessary to achieve a high level of deuteration at  $\beta$  positions, which in turn is a prerequisite for recording artifact-free  $^1\text{H}\alpha$ -relaxation dispersion profiles. Dispersion curves for 17 residue-types (the exceptions are Gly, Ser and Thr) can be measured from one experiment<sup>44</sup>, whereas Gly requires a modified scheme<sup>45</sup>. Ser and Thr are not amenable for analysis using this labeling approach and the current generation of published NMR experiments. An elegant alternative way of achieving protonation at the  $\alpha$  position in an otherwise deuterated background has been described by Yamazaki *et al.*<sup>51</sup>. This method is based on racemization accompanied by the acetylation of  $^{13}\text{C}/^{15}\text{N}/^2\text{H}$  labeled amino acids produced by the hydrolysis of triple-labeled proteins. For those residues types that are amenable to this type of chemistry, impressive levels of protonation at the  $\alpha$  position are achieved ( $\approx 80\%$ ), leading to spectra of better sensitivity when compared with our approach; however, the method does not work for almost half of the residue types. Our scheme is thus more general and in addition is significantly simpler to implement.

In all, 19 out of the 20 residue-types are enriched at  $\text{C}\alpha$  in Sample 3 with Leu being the exception<sup>46</sup>. However, both Ile and Val contain  $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$  isotopomers and as such are not amenable to study using existing dispersion experiments<sup>31</sup>. The excited state  $^{13}\text{C}\alpha$  shifts for these residues in some cases can, however, be obtained using slight modifications to the published experiments (unpublished data; experiments available upon request from the authors). Although Gly residues are  $^{13}\text{C}$  labeled at the  $\text{C}\alpha$  position and this label is isolated, we prefer to use Sample 2 for measurement of  $^{13}\text{C}\alpha$  dispersions for this residue, as in this case the experiment is carried out on a  $^{13}\text{CHD}$  spin system, which reduces the intrinsic transverse relaxation rate compared with the case with  $^{13}\text{CH}_2$  groups<sup>45</sup>. The fractional incorporation of  $^{13}\text{C}$  in the  $\text{C}\alpha$  position ranges from 17 to 45% for all residue-types except Leu, and was determined by comparing signal intensities in  $^{13}\text{C}$ - $^1\text{H}$  spectra recorded on this sample and on a uniformly  $^{13}\text{C}/^{15}\text{N}$  labeled sample after normalization with the average intensity in  $^{15}\text{N}$ - $^1\text{H}$  spectra to account for different sample concentrations, **Table 3**.

The two remaining samples (Samples 4 and 5 of **Table 1**) are produced for recording  $^{13}\text{C}\beta$ -relaxation dispersion experiments. In the case of Sample 4 the labeling protocol generates  $^{13}\text{C}\beta$  probes that are well isolated for 11 residue-types (the fraction of molecules with  $^{12}\text{C}\alpha$ - $^{13}\text{C}\beta$ - $^{12}\text{C}\gamma$  is over 93%), leading to artifact-free dispersion experiments for these amino acids, with fractional incorporation of label in the range 33-43%, determined in the same way as previously described for Sample 3, **Table 4**. An additional four residue-types are available from Sample 5, with incorporation of  $^{13}\text{C}$  label at the  $\beta$  position varying from 29 to 39%, as indicated in **Table 4**. Ile, Leu and Val are not amenable for analysis using this labeling procedure because they are not enriched at  $\text{C}\beta$  in Sample 4, whereas for Sample 5 a large fraction of the  $^{13}\text{C}\beta$  probes are present as  $^{13}\text{C}\beta$ - $^{13}\text{C}\alpha/\gamma$  spin-pairs.

**TABLE 2** | Predicted overall deuteration level at the  $\beta$ -positions in a protein expressed with uniformly deuterated glucose as the carbon source in 50%  $\text{D}_2\text{O}/50\%$   $\text{H}_2\text{O}$  (ref. 44).

Amino acid	Overall deuteration
Cys, Phe, Ser, Trp and Tyr	0.88
Arg, Gln, Glu, Leu and Pro	0.50
Asn, Asp and Met	0.66
Ile and Val	0.50
Ala	0.75
His	0.75
Lys	0.70
Thr	0.66

**TABLE 3** | Fractional <sup>13</sup>C enrichment at C $\alpha$  for different amino acid residue-types produced by overexpression in BL21(DE3) cells using [2-<sup>13</sup>C]-glucose as the carbon source<sup>a</sup> (ref. 46).

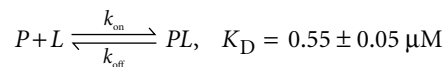
Residues	Precursor(s)	Enrichment at C $\alpha$
Ala, Cys, Gly, Phe, Ser, Trp, Tyr and Val	Pyruvate, 3-phosphoglycerate, phosphoenolpyruvate	0.45 ± 0.04 (23) <sup>b</sup>
Asn, Asp, Ile, Met and Thr	Oxaloacetate	0.28 ± 0.02 (21) <sup>b</sup>
Arg, Gln, Glu and Pro	$\alpha$ -ketoglutarate	0.17 ± 0.02 (17)
Lys	Oxaloacetate or pyruvate	0.38 ± 0.03 (11)
His	Ribose-5-phosphate	0.30 ± n.a. (1)
Leu	Acetyl-S-CoA	0.06 ± 0.01 (7)

<sup>a</sup>C $\alpha$ -enrichment levels are based on data from ubiquitin and the FBP11 FF domain. Data for residue-types derived from similar precursors have been averaged. <sup>b</sup>For Ile and Val both <sup>13</sup>C $\alpha$ -<sup>13</sup>C $\beta$  and <sup>13</sup>C $\alpha$ -<sup>13</sup>C $\beta$  moieties are observed in relative proportions of 1.2:1.

As described above, the use of [1-<sup>13</sup>C]- and [2-<sup>13</sup>C]-glucose in the production of Samples 3–5 ensures that optimal protein yields are obtained. A disadvantage is that the maximum fractional <sup>13</sup>C incorporation at C $\alpha$  or C $\beta$  is only 50%, as these positions are equally likely to be derived from the unlabeled 5 or 6 positions as from the corresponding labeled 2 or 1 positions of glucose. Use of [1,6-<sup>13</sup>C<sub>2</sub>]- and [2,5-<sup>13</sup>C<sub>2</sub>]-glucose would double the theoretical maximum fractional incorporation, whereas preserving the advantages of using glucose as the carbon source, but unfortunately these compounds are restrictively expensive. There are, however, other, more affordable carbon sources that have advantages in terms of incorporation levels, although maybe at the expense of lower levels of protein overexpression. One of these is glycerol labeled at either position 2 or positions 1 and 3 (ref. 36), which is frequently used to produce proteins for solid state NMR applications<sup>52,53</sup>. Especially for Samples 3 and 5 it may be worth considering substituting glycerol for glucose, as the protein yields using the two carbon sources are very similar even for the BL21(DE3) $\Delta$ sdh strain used for Sample 5 (unpublished results), and the prices of [2-<sup>13</sup>C]-glucose and [2-<sup>13</sup>C]-glycerol do not differ much. We have not analyzed the fraction of isolated <sup>13</sup>C $\alpha$  and <sup>13</sup>C $\beta$  moieties for samples using [2-<sup>13</sup>C]-glycerol as the carbon source, but the work by LeMaster and Kushlan<sup>36</sup> suggests similar values to those we have obtained using [2-<sup>13</sup>C]-glucose. For Sample 4 [1-<sup>13</sup>C]-glucose is, however, preferable, as [1,3-<sup>13</sup>C<sub>2</sub>]-glycerol is more than four times as expensive, offsetting any advantages in terms of increased fractional incorporation. An alternative to glycerol for Sample 3 is pyruvate labeled at position 2. However, although protein yields for pyruvate and glucose are almost identical, we have noticed significantly more scrambling, leading to deleterious <sup>13</sup>C-<sup>13</sup>C spin pairs in the case of pyruvate, at least if protein expression is carried out overnight<sup>31,37,43</sup>. Furthermore, pyruvate is not an option for Samples 4 and 5, as the BL21(DE3) $\Delta$ sdh strain does not grow on this carbon source (unpublished results).

### Validation of methodology

The ultimate test of efficacy of this proposed labeling strategy is the accuracy of the excited state chemical shifts that are extracted from CPMG-relaxation dispersion experiments. This can be established using an exchanging system



where *P* is the 7-kDa SH3 domain from the yeast protein Abp1p<sup>54–56</sup>, *L* is a 17 residue peptide from the protein Ark1p<sup>57</sup> and *PL* is the complex formed between the two. Addition of 5% mole fraction *L* leads to an exchange rate,  $k_{\text{ex}} = k_{\text{on}}[L] + k_{\text{off}} \approx 300 \text{ s}^{-1}$  at 25 °C<sup>34</sup>, well within the window of quantification by CPMG-relaxation dispersion methodology. The addition of a small mole fraction of *L* (for example 5%) renders *P* the visible ground state (95% populated), whereas *PL* is invisible (5% populated, resonances significantly broadened because of exchange). The chemical shifts of *PL* can be determined from CPMG experiments and

**TABLE 4** | Fractional <sup>13</sup>C enrichment at C $\beta$  for different amino acid residue-types produced by overexpression in BL21(DE3) $\Delta$ sdh cells using the indicated carbon sources<sup>a</sup> (ref. 38).

Amino acid	[1- <sup>13</sup> C]-glucose/ NaH <sup>12</sup> CO <sub>3</sub>	[2- <sup>13</sup> C]-glucose
Ala	<b>0.41 ± 0.002 (2)</b> <sup>a</sup>	0.09 ± 0.008(2)
Arg <sup>b</sup>	n.a.	<b>n.a.</b>
Asn	<b>0.33 ± 0.04 (3)</b>	0.11 ± 0.02 (3)
Asp	<b>0.35 ± 0.007 (2)</b>	0.12 ± 0.02 (2)
Cys <sup>b</sup>	<b>n.a.</b>	n.a.
Gln <sup>b</sup>	n.a.	<b>n.a.</b>
Glu	0.05 ± 0.01 (2)	<b>0.39 ± 0.05 (2)</b>
His <sup>b</sup>	n.a.	n.a.
Ile	n.a.	0.43 ± 0.04 (3)
Leu	n.a.	0.42 ± 0.04 (4)
Lys	<b>0.33 ± 0.05 (3)</b>	0.11 ± 0.007 (3)
Met <sup>c</sup>	<b>n.a.</b>	n.a.
Phe	<b>0.33 ± 0.0004 (2)</b>	0.04 ± 0.0008 (2)
Pro	0.05 ± 0.006 (2)	<b>0.29 ± 0.01 (2)</b>
Ser	<b>0.43 ± 0.08 (3)</b>	0.12 ± 0.01 (3)
Thr	<b>0.34 ± 0.002 (2)</b>	0.10 ± 0.009 (2)
Trp	<b>0.35 ± 0.03 (2)</b>	0.08 ± 0.02 (4)
Tyr	<b>0.36 ± 0.02 (2)</b>	0.06 ± 0.003 (2)
Val	n.a.	0.46 ± n.a. (1)

n.a., not applicable.

<sup>a</sup>C $\beta$ -enrichment levels are based on data from the Abp1p SH3 domain. Residues highlighted in bold are those that can be used for measurement of <sup>13</sup>C $\beta$ -relaxation dispersion profiles, with the labeling scheme described (i.e., sufficient levels of enrichment and >93% isolation). <sup>b</sup>These residue-types are not present in the Abp1p SH3 domain. Here, bold indicates that they are expected to be enriched at C $\beta$  based on an analysis of metabolic pathways<sup>50</sup>. <sup>c</sup>The <sup>13</sup>C enrichment at C $\beta$  could not be quantified because of a spectral overlap in the spectrum of the uniformly <sup>13</sup>C-labeled protein; see reference 38 for details.

subsequently compared with those measured directly from peak positions in spectra recorded on samples for which saturating amounts of *L* have been added so that *PL* becomes the visible, ground state. As shown in **Table 5**, the pair-wise root mean square deviations (RMSDs) between CPMG and direct measures of chemical shifts are <0.05 ppm for all nuclei<sup>31,43,44</sup> with the exception of <sup>13</sup>Cβ, wherein it is 0.16 ppm and 0.27 ppm from measurements recorded on Samples 4 and 5, respectively<sup>38</sup>. All of these values are significantly better than the RMSDs between predicted and observed chemical shifts estimated using the best current prediction algorithm, shift prediction from analogy in residue type and torsion angle (SPARTA<sup>32</sup>).

Experiments involving Samples 1–3 have so far been successfully applied to the 18.8-kDa T4 lysozyme and in the case of experiments involving Sample 1 to the 21.6-kDa dimer of HIV protease<sup>42,58</sup>. Applications of <sup>15</sup>N-relaxation dispersion spectroscopy to proteins with molecular weights on the order of 50 kDa have also been reported, making use of TROSY (transverse relaxation optimized spectroscopy)-based experiments in these cases<sup>39</sup>. The <sup>13</sup>Cβ-relaxation dispersion experiments involving Samples 4 and 5 are likely limited to proteins of molecular weight <10 kDa because of problems with resonance overlap in the <sup>13</sup>Cβ–<sup>1</sup>Hβ region of correlation spectra. In this regard the use of Samples 4 and 5, with signals from a number of amino acid residue-types either missing or significantly attenuated in each of the two samples, **Table 4**, is advantageous. A second limitation to this methodology is that relatively large protein concentrations (≈1 mM) are required for some of these experiments, as the isotopic enrichment at the relevant positions in many cases is not complete.

In summary we have described labeling methodology for the production of five samples that can be used for the measurement of <sup>13</sup>Cα, <sup>13</sup>Cβ, <sup>13</sup>CO, <sup>1</sup>Hα, <sup>1</sup>H<sup>N</sup> and <sup>15</sup>N chemical shifts of invisible protein states that exchange with visible, ground conformers on the millisecond time-scale. This is accomplished by fitting measured CPMG-relaxation dispersion profiles that are recorded using NMR experiments that have been designed to exploit the labeling patterns that are produced (**Table 1**). The labeling schemes have been shown to be general in the sense that they produce the desired label for the majority of the residue-types in a protein, they give high levels of protein overexpression and fractional incorporation of label and they are easy to implement. The utility of the approach is demonstrated by the fact that accurate excited state chemical shifts can be extracted in CPMG experiments involving these samples. This methodology thus provides a very robust and powerful approach for studies of excited protein conformers that often have critical roles in biological function.

**ACKNOWLEDGMENTS** This work was supported by funding from the Canadian Institutes of Health Research (CIHR) to L.E.K. P.L. is supported by the Swedish Research Council and D.F.H. holds a CIHR postdoctoral fellowship. L.E.K. is the recipient of a Canada Research Chair in Biochemistry.

**AUTHOR CONTRIBUTIONS** P.L. and L.E.K. conceived the labeling strategies. P.L., P.V. and D.F.H. produced the protein samples and validated the methodology. P.L. and L.E.K. wrote the paper. L.E.K. supervised the project.

Published online at <http://www.natureprotocols.com/>.  
Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

- Mittermaier, A. & Kay, L.E. New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224–228 (2006).
- Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
- Goto, N.K. & Kay, L.E. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.* **10**, 585–592 (2000).
- Sattler, M. *et al.* Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Reson. Spectrosc.* **34**, 93–158 (1999).
- Carr, H.Y. & Purcell, E.M. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* **94**, 630–638 (1954).
- Meiboom, S. & Gill, D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* **29**, 688–691 (1958).

**TABLE 5** | RMSD between  $|\Delta\sigma|$  obtained from Carr–Purcell–Meiboom–Gill (CPMG) experiments and direct measurements of peak positions, where  $\Delta\sigma$  is the difference in chemical shifts between ground and excited states, in p.p.m.

Nucleus	Labeling scheme	RMSD <sub>CPMG,DIRECT</sub>
<sup>15</sup> N	Sample 1	0.030 p.p.m.
<sup>1</sup> H <sup>N</sup>	Sample 1	0.013 p.p.m.
<sup>13</sup> CO	Sample 1	0.024 p.p.m.
<sup>1</sup> Hα	Sample 2	0.022 p.p.m.
<sup>13</sup> Cα	Sample 3	0.040 p.p.m.
<sup>13</sup> Cβ	Sample 4	0.16 p.p.m.
<sup>13</sup> Cβ	Sample 5	0.27 p.p.m.

- Hill, R.B. *et al.* Molecular motions and protein folding: characterization of the backbone dynamics and folding equilibrium of α<sub>2</sub>D using <sup>13</sup>C NMR spin relaxation. *J. Am. Chem. Soc.* **122**, 11610–11619 (2000).
- Korzhnev, D.M. *et al.* Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586–590 (2004).
- Zeeb, M. & Balbach, J. NMR spectroscopic characterization of millisecond protein folding by transverse relaxation dispersion measurements. *J. Am. Chem. Soc.* **127**, 13207–13212 (2005).
- Sugase, K. *et al.* Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
- Eisenmesser, E.Z. *et al.* Enzyme dynamics during catalysis. *Science* **295**, 1520–1523 (2002).
- Wolf-Watz, M. *et al.* Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat. Struct. Mol. Biol.* **11**, 945–949 (2004).
- Eisenmesser, E.Z. *et al.* Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **438**, 117–121 (2005).
- Boehr, D.D. *et al.* The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642 (2006).
- Vallurupalli, P. & Kay, L.E. Complementarity of ensemble and single-molecule measures of protein motion: a relaxation dispersion NMR study of an enzyme complex. *Proc. Natl. Acad. Sci. USA* **103**, 11910–11915 (2006).
- Watt, E.D. *et al.* The mechanism of rate-limiting motions in enzyme function. *Proc. Natl. Acad. Sci. USA* **104**, 11981–11986 (2007).
- Mulder, F.A.A. *et al.* Studying excited states of proteins by NMR spectroscopy. *Nat. Struct. Biol.* **8**, 932–935 (2001).
- Palmer A.G. III *et al.* Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **339**, 204–238 (2001).



19. Skrynnikov, N.R. *et al.* Reconstructing NMR spectra of 'invisible' excited protein states using HSQC and HMQC experiments. *J. Am. Chem. Soc.* **124**, 12352–12360 (2002).
20. Korzhnev, D.M. & Kay, L.E. Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. *Acc. Chem. Res.* **41**, 442–451 (2008).
21. Spera, S. & Bax, A. Empirical correlation between protein backbone conformation and C $\alpha$  and C $\beta$  <sup>13</sup>C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490–5492 (1991).
22. Wishart, D.S. & Sykes, B.D. The <sup>13</sup>C chemical-shift index—a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *J. Biomol. NMR* **4**, 171–180 (1994).
23. Wishart, D.S. & Case, D.A. Use of chemical shifts in macromolecular structure determination. *Methods Enzymol.* **338**, 3–34 (2002).
24. Shen, Y. *et al.* Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. USA* **105**, 4685–4690 (2008).
25. Cavalli, A. *et al.* Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. USA* **104**, 9615–9620 (2007).
26. Tjandra, N. & Bax, A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278**, 1111–1114 (1997).
27. Tolman, J.R. *et al.* Nuclear magnetic dipole interactions in field-oriented proteins—information for structure determination in solution. *Proc. Natl. Acad. Sci. USA* **92**, 9279–9283 (1995).
28. Kay, L.E. *et al.* 3-Dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **89**, 496–514 (1990).
29. Bax, A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci.* **12**, 1–16 (2003).
30. Prestegard, J.H. *et al.* Determination of protein backbone structures from residual dipolar couplings. *Methods Enzymol.* **394**, 175–209 (2005).
31. Hansen, D.F. *et al.* Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy. How well can we do? *J. Am. Chem. Soc.* **130**, 2667–2675 (2008).
32. Shen, Y. & Bax, A. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR* **38**, 289–302 (2007).
33. Igumenova, T.I. *et al.* Characterization of chemical exchange using residual dipolar coupling. *J. Am. Chem. Soc.* **129**, 13396–13397 (2007).
34. Vallurupalli, P. *et al.* Measurement of bond vector orientations in invisible excited states of proteins. *Proc. Natl. Acad. Sci. USA* **104**, 18473–18477 (2007).
35. Kay, L.E. *et al.* Backbone dynamics of proteins as studied by <sup>15</sup>N inverse detected heteronuclear NMR-spectroscopy—application to Staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).
36. LeMaster, D.M. & Kushlan, D.M. Dynamical mapping of *E. coli* thioredoxin via <sup>13</sup>C NMR relaxation analysis. *J. Am. Chem. Soc.* **118**, 9255–9264 (1996).
37. Mulder, F.A.A. *et al.* Slow internal dynamics in proteins: application of NMR relaxation dispersion spectroscopy to methyl groups in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.* **124**, 1443–1451 (2002).
38. Lundström, P. *et al.* Measuring <sup>13</sup>C chemical shifts of invisible excited states in proteins by relaxation dispersion NMR spectroscopy. *J. Biomol. NMR* **44**, 139–155 (2009).
39. Loria, J.P. *et al.* A relaxation-compensated Carr–Purcell–Meiboom–Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332 (1999).
40. Tollinger, M. *et al.* Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* **123**, 11341–11352 (2001).
41. Ishima, R. & Torchia, D.A. Extending the range of amide proton relaxation dispersion experiments in proteins using a constant-time relaxation-compensated CPMG approach. *J. Biomol. NMR* **25**, 243–248 (2003).
42. Ishima, R. *et al.* Carbonyl carbon transverse relaxation dispersion measurements and ms– $\mu$ s timescale motion in a protein hydrogen bond network. *J. Biomol. NMR* **29**, 187–198 (2004).
43. Lundström, P. *et al.* Measurement of carbonyl chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy: comparison between uniformly and selectively <sup>13</sup>C labeled samples. *J. Biomol. NMR* **42**, 35–47 (2008).
44. Lundström, P. *et al.* Accurate measurement of alpha proton chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 1915–1926 (2009).
45. Vallurupalli, P. *et al.* CPMG relaxation dispersion NMR experiments measuring glycine <sup>1</sup>H $\alpha$  and <sup>13</sup>C $\alpha$  chemical shifts in the 'invisible' excited states of proteins. *J. Biomol. NMR* **45**, 45–55 (2009).
46. Lundström, P. *et al.* Fractional <sup>13</sup>C enrichment of isolated carbons using [1-<sup>13</sup>C]- or [2-<sup>13</sup>C]-glucose facilitates the accurate measurement of dynamics at backbone C $\alpha$  and side-chain methyl positions in proteins. *J. Biomol. NMR* **38**, 199–212 (2007).
47. Tsumoto, K. *et al.* Practical considerations in refolding proteins from inclusion bodies. *Prot. Expr. Pur.* **28**, 1–8 (2003).
48. Tsumoto, K. *et al.* Review: why is arginine effective in suppressing aggregation? *Protein Peptide Lett.* **12**, 613–619 (2005).
49. Middelberg, A.P.J. Preparative protein refolding. *Trends Biotechnol.* **20**, 437–443 (2002).
50. Voet, D. & Voet, J.G. *Biochemistry* 3rd edn. 1030–1045 (John Wiley & Sons, Inc., Hoboken, NJ, 1995).
51. Yamazaki, T. *et al.* Assignment of backbone resonances for larger proteins using the C-13-H-1 coherence of a H-1(alpha)-, H-2-, C-13-, and N-15-labeled sample. *J. Am. Chem. Soc.* **119**, 872–880 (1997).
52. Castellani, F. *et al.* Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 98–102 (2002).
53. Wasmer, C. *et al.* Amyloid fibrils of the HET-s(218–289) prion form a beta solenoid with a triangular hydrophobic core. *Science* **319**, 1523–1526 (2008).
54. Rath, A. & Davidson, A.R. The design of a hyperstable mutant of the Abp1p SH3 domain by sequence alignment analysis. *Protein Sci.* **9**, 2457–2469 (2000).
55. Drubin, D.G. *et al.* Homology of a yeast actin-binding protein to signal transduction proteins and myosin-I. *Nature* **343**, 288–290 (1990).
56. Lila, T. & Drubin, D.G. Evidence for physical and functional interactions among two *Saccharomyces cerevisiae* SH3 domain proteins, an adenyl cyclase-associated protein and the actin cytoskeleton. *Mol. Biol. Cell* **8**, 367–385 (1997).
57. Haynes, J. *et al.* The biologically relevant targets and binding affinity requirements for the function of the yeast actin-binding protein 1 Src-homology 3 domain vary with genetic context. *Genetics* **176**, 193–208 (2007).
58. Ishima, R. *et al.* Using amide <sup>1</sup>H and <sup>15</sup>N transverse relaxation to detect millisecond time-scale motions in perdeuterated proteins: application to HIV-1 protease. *J. Am. Chem. Soc.* **120**, 10534–10542 (1998).