Journal of Biomolecular NMR **29:** 289–297, 2004.

Uniform and residue-specific ^{15}N-labeling of proteins on a highly deuterated background

Jocelyne Fiaux[a,**], Eric B. Bertelsen[b,**], Arthur L. Horwich[b] & Kurt Wüthrich[a,*]
[a]*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Zürich, CH–8093 Zürich, Switzerland;* [b]*Howard Hughes Medical Institute and Department of Genetics, Yale School of Medicine, New Haven, CT 06510, U.S.A.*

Received 20 November 2003; Accepted 21 January 2004

Key words: large molecular structures, residue-specific isotope labeling, solution NMR, stable-isotope labeling, uniform deuterium labeling

Abstract

A general method for stable-isotope labeling of large proteins is introduced and applied for studies of the *E. coli* GroE chaperone proteins by solution NMR. In addition to enabling the residue-specific ^{15}N-labeling of proteins on a highly deuterated background, it is also an efficient approach for uniform labeling. The method meets the requirements of high-level deuteration, minimal cross-labeling and high protein yield, which are crucial for NMR studies of structures with sizes above 150 kDa. The results obtained with the new protocol are compared to other strategies for protein labeling, and evaluated with regard to the influence of external factors on the resulting isotope labeling patterns. Applications with the GroE system show that these strategies are efficient tools for studies of structure, dynamics and intermolecular interactions in large supramolecular complexes, when combined with TROSY- and CRINEPT-based experimental NMR schemes.

Introduction

Studies of large macromolecular structures by solution NMR are complicated by a number of factors, including line-broadening caused by rapid transverse spin relaxation, signal overlap due to the high number of resonances, and low sensitivity when only limited concentrations can be achieved. TROSY- and CRINEPT-based NMR techniques can suppress fast transverse relaxation and enable studies of structures with sizes of 100 kDa and above (Pervushin et al., 1997; Riek et al., 1999; Salzmann et al., 2000), and the sensitivity has been enhanced with the introduction of higher-field instruments and cryo-probe technology. However, these advances in NMR spectroscopy can only be fully exploited in combination with suitable stable-isotope labeling strategies, in particular the replacement of non-exchangeable protons for deuterons, and residue-specific ^{15}N- and/or ^{13}C labeling. Different protocols to achieve high levels of deuteration have been reported (Gardner and Kay, 1998; Leiting et al., 1998; Lian and Middleton, 2001; Marley et al., 2001), but most of the currently used techniques depend upon lengthy conditioning of the cells to the deuterated growth medium, which is sometimes not well tolerated by the expression system. Furthermore, although residue-specific labeling has quite routinely been accomplished in a protonated background (McIntosh and Dahlquist, 1990; Muchmore et al., 1989), specific labeling on a deuterated background typically involves the addition of protonated ^{15}N- or ^{15}N,^{13}C-labeled amino acids to the expression medium (Kelly et al., 1999; Metzler et al, 1996; Smith et al., 1996). These protocols are not suitable for structure sizes above 100 kDa, because the residual protonation of the polypeptide chain deteriorates the NMR spectra due to fast relaxation.

Here we describe a procedure for achieving residue-specific ^{15}N-labeling of proteins on a highly

*To whom correspondence should be addressed. E-mail: wuthrich@mol.biol.ethz.ch
**These authors contributed equally to this work.

deuterated background. It is presented in the form of a general protocol, which we have applied for both uniform and residue-specific labeling. We describe several variants of this protocol and discuss the different factors affecting the resulting labeling pattern. Practical use of the protocol is illustrated with proteins of the *E. coli* GroE system, with structure sizes from 70 to 870 kDa. The discussion and comparison of different labeling strategies in this publication complement previous reports on experimental NMR schemes for work with large structures (Fiaux et al., 2002; Riek et al., 2002).

Materials and methods

Materials

Deuterated, or ^{15}N-enriched and deuterated algal hydrolysate (CELTONE-d or CELTONE-dN base powder), and purified [^{15}N,^{2}H]-lysine and [^{15}N,^{2}H]-cysteine were purchased from Spectra Stable Isotopes, Inc., Columbia, MD. [^{15}N,^{2}H]-valine, [^{15}N,^{2}H]-leucine, and $^{15}NH_4Cl$ were obtained from Cambridge Isotope Laboratories, Inc., Andover, MA. 99.8% $^{2}H_2O$ (D_2O) was purchased from Isotec, Miamisburg, OH.

The strains DL39 (F-, LAM-, aspC13, fnr-25, rph-1, ilvE12, tyrB507) and JM15 (F-, cysE50, tfr-8) were obtained from the *E. coli* Genetic Stock Center, Yale University, and modified to contain the T7 expression system, yielding DL39(DE3) and JM15(DE3), respectively. The strains CT8 and CT19 obtained from Hoffmann-LaRoche, Inc., Basel, Switzerland are derivatives of BL21(DE3), which carry the mutation lysA23, and the mutations aspC, ilvE, trpB and tyrB, respectively (Waugh, 1996). The genetic lesions in these strains affect the following enzymes: aspC, aspartate aminotransferase; cysE, serine acetyltransferase; ilvE, branched chain amino acid aminotransferase; lysA, diaminopimelate decarboxylase; trpB, tryptophan synthase; tyrB, aromatic amino acid aminotransferase.

Protein expression

Recombinant GroEL, GroES and human dihydrofolate reductase (hDHFR) were expressed from pET expression vectors. Uniformly ^{15}N-labeled and deuterated proteins were expressed in the BL21(DE3) strain. For residue-specific ^{15}N-labeling of valine and leucine, the proteins were expressed in the transaminases-deficient strains DL39(DE3) or CT19 (LeMaster and Richards, 1988; Waugh, 1996). For labeling of lysine and cysteine, the proteins were expressed either from the corresponding auxotrophic cell lines (Waugh, 1996) obtained from Hoffmann-LaRoche, Inc., Basel, Switzerland, or from the BL21(DE3) strain. Table 3 indicates the strain type used for each protein batch.

Uniformly or residue-specifically ^{15}N-labeled and deuterated proteins were expressed in cells grown on a CELTONE-supplemented medium (Table 1), and a variety of isotope-labeled compounds were used in various combinations to achieve different labeling patterns (Tables 2 and 3). In all cases, the cultures were inoculated from a preculture grown in the same medium, and protein expression was induced with

Table 1. Composition of the CELTONE-supplemented media used in the present study[a]

Basic minimal medium[b]		
→	800 mL	H_2O or D_2O
	100 mL	M9 salt solution[c]
	2 mL	1M $MgSO_4$
→	1 g	NH_4Cl
→	1 g	D-glucose
Vitamin and trace element supplements (optional)		
	10 mL	Vitamin mix[d]
	2 mL	Trace element solution[e]
Amino acid supplements		
→	1–3 g	deuterated algal lysate amino acid mixture (CELTONE base powder), dissolved at 30 mg/mL and filtered[f].
		sterile filter
		add antibiotics
→		for specific amino acid labeling, add the [^{15}N,^{2}H]-amino acid in proper amount (see text) at induction

[a]The arrows indicate the compounds that may be isotope-labeled in the different variants of the protocol (see text).
[b]Based on Sambrook et al. (1989).
[c]$Na_2HPO_4 \cdot 2H_2O$ 85.1 g/L, KH_2PO_4 30 g/L, NaCl 5 g/L.
[d]Thiamine 0.5 mg/mL, d-biotin 0.1 mg/mL, Choline chloride 0.1 mg/mL, Folic acid 0.1 mg/mL, Niacinamide 0.1 mg/mL, D-pantothenic acid 0.1 mg/mL, Pyridoxal 0.1 mg/mL, Riboflavin 0.01 mg/mL.
[e]5M HCl 8 mL/L, $FeCl_2 \cdot 4H_2O$ 5 g/L, $CaCl_2 \cdot 2H_2O$ 184 mg/L, H_3BO_3 64 mg/L, $CoCl_2 \cdot 6H_2O$ 18 mg/L, $CuCl_2 \cdot 2H_2O$ 4 mg/L, $ZnCl_2$ 340 mg/L, $Na_2MoO_4 \cdot 2H_2O$ 605 mg/L, $MnCl_2 \cdot 4H_2O$ 40 mg/L.
[f]Can also be added just before induction for pulse-labeling.

Table 2. Procedures used for the preparation of deuterated, uniformly ^{15}N-labeled proteins

	Medium description	Deuteration	Yield/L culture	Advantages/disadvantages
1	Minimal medium on unlabeled glucose, D_2O, $^{15}NH_4Cl$[a]	∼85%	GroES: 65 mg SR1: 50 mg	+low cost, easily performed −residual protonation in side chains
2	Minimal medium on [$^2H_4$]-acetate, D_2O, $^{15}NH_4Cl$[b]	∼98%	GroEL: 24 mg GroES: 20 mg	+high level of deuteration −slow growth, low yield
3	Martek-9dN on D_2O[c]	∼98%	SR1: 20 mg	−low yield with BL21 cells, expensive
4	Minimal medium on glucose supplemented with CELTONE-dN in H_2O[d]	– (∼30–50% in pulse labeling)[e]	GroES: 38 mg[b] GroEL: 50 mg[b] SR1: 35 mg[b]	+no N–H/N–D exchange problems

[a]Leiting et al. (1998); Shekhtman et al. (2002).
[b]Venters et al. (1996).
[c]Spectra Stable Isotopes, Columbia, MD. This medium corresponds to a minimal medium in D_2O supplemented with 1 g/L algal amino acid mix CELTONE-dN (see Table 1).
[d]See also Löhr et al. (2003), Markus et al. (1994), Serber et al. (2001) and Yamazaki et al. (1997).
[e]Pulse labeling: labeled CELTONE added at induction.

0.5–1 mM IPTG in the mid-exponential growth phase. The cells were collected after 3 to 5 h.

For comparison, uniformly ^{15}N-labeled and deuterated proteins were also produced using standard minimal-medium protocols in D_2O (Leiting et al., 1998; Sambrook et al., 1989; Shekhtman et al., 2002; Venters et al., 1996), and residue-specifically ^{15}N-labeled and deuterated proteins were prepared using variants of published protocols, as described in the Tables 2 and 3.

Protein preparation and characterization

GroEL (Fiaux et al., 2002; Weissman et al., 1995), SR1 (Horwich et al., 1998), GroES (Fiaux et al., 2002; Weissman et al., 1995), and hDHFR (Goldberg et al., 1997) were extracted and purified as described previously.

The GroES–SR1 complexes were prepared by addition of 4.2 mg of isotope-labeled GroES to 34 mg natural isotope abundance SR1 in the presence of 10 mM ATP.

The overall levels of deuteration, protonation at the C^α position, and ^{15}N cross-labeling were assessed by a combination of 1D ^{1}H-NMR spectroscopy and [^{15}N,^{1}H]-HSQC spectroscopy of the protein under denaturing conditions (see below). Deuteration was further quantified by Matrix Assisted Laser Desorption Ionization Time of Flight Mass Spectrometry (MALDI-TOF-MS) on a Voyager-DE Elite Biospectrometry workstation (PerSeptive Biosystems, Applied Biosystems). The levels of isotope labeling were deduced from comparison of the measured molecular mass of the labeled protein to that of the unlabeled reference sample of the same protein.

NMR measurements

For studies of the folded proteins, NMR samples were prepared in H_2O solutions containing 25 mM potassium phosphate at pH 6.15, 20 mM KCl, 5% D_2O, and 0.02% NaN_3. For measurements in the denatured state, the proteins were dissolved in 6M urea or 6M guanidinium chloride solutions containing 25 mM potassium phosphate at pH 6.15, 20 mM KCl, 5% D_2O, and 0.02% NaN_3. The NMR experiments were carried out at a ^{1}H resonance frequency of 750 MHz or 900 MHz on Bruker DRX 750 and DRX 900 spectrometers. The sample temperature was 25 °C. 2D [^{15}N,^{1}H]-HSQC, 2D [^{15}N,^{1}H]-TROSY and 2D [^{15}N,^{1}H]-CRIPT-TROSY experiments were recorded as described previously (Fiaux et al., 2002; Riek et al., 2002). The spectra were processed with the program PROSA (Güntert et al., 1992) and analyzed with the program XEASY (Bartels et al., 1995).

Results and discussion

In this study we searched for a general scheme for isotope labeling of proteins in large molecular structures to address specific requirements arising with the use of TROSY- and CRINEPT-based experiments (Pervushin et al., 1997, 2000; Riek et al., 2000, 2002;

Table 3. Procedures used for the preparation of deuterated and residue-specifically ^{15}N-labeled proteins

	Medium description	Deuteration	Yield/L culture	*E. coli* strain used	Advantages/disadvantages
1	Protonated synthetic rich medium in H_2O +[^{15}N,^{2}H]-a.a.[a]	specific, only 1 a.a.	[^{15}N,^{2}H-Leu]-GroEL:126 mg	BL21(DE3)	+no N–H/N–D exchange problems, high yield, low cost −deuteration not sufficient for NMR with large molecules
2	Minimal medium on glucose+ CELTONE-d + [^{15}N,^{2}H]-a.a.				
	in H_2O:	∼60–92%	[^{15}N,^{2}H-Leu, $U-^{2}$H]-GroEL: 20 mg	DL39(DE3)[c]	+no N–H/N–D exchange problems
			[^{15}N,^{2}H-Leu, $U-^{2}$H]-SR1: 5 mg	DL39(DE3)	+intermediate deuteration can be achieved
			[^{15}N,^{2}H-Leu, $U-^{2}$H]-GroES: 50 mg	CT19	
	in D_2O:	∼95–97%	[^{15}N,^{2}H-Val, $U-^{2}$H]-GroES: 30 mg	CT19	+high deuteration
			[^{15}N,^{2}H-Lys, $U-^{2}$H]-GroES: 80 mg	CT8[d]	
			[^{15}N,^{2}H-Cys, $U-^{2}$H]-DHFR: 3 mg	JM15(DE3)[c]	
3	Martek-9d in D_2O + [^{15}N,^{2}H]-a.a.	∼97%	[^{15}N,^{2}H-Leu, U-^{2}H]-SR1: 15 mg	BL21(DE3)[e]	
4	LB–> minimal medium[f] on glucose + CELTONE-d + [^{15}N,^{2}H]-a.a.				+improved growth
	in H_2O:	∼84%	[^{15}N,^{2}H-Leu, $U-^{2}$H]-DHFR: 4 mg	DL39(DE3)	
	in D_2O:	∼93%	[^{15}N,^{2}H-Leu, $U-^{2}$H]-DHFR: 4 mg	DL39(DE3)	

[a]Muchmore and Dahlquist (1989).
[b]The deuteration level depends on the amount of amino acid in the medium at the induction time point (see text).
[c]Obtained from the E. coli Genetic Stock Center, Yale University, and modified as described in Methods.
[d]Waugh (1996). Obtained from Hoffmann-LaRoche, Basel. CT8 and CT19 are BL21(DE3) derivatives carrying the lysA23 mutation and the aspC, ilvE, trpB and tyrB mutations, respectively. The defective genes in these strains correspond to the following enzymes: aspC, aspartate aminotransferase; cysE, serine acetyltransferase; ilvE, branched chain amino acid aminotransferase; lysA, diaminopimelate decarboxylase; trpB, tryptophan synthase; tyrB, aromatic amino acid aminotransferase.
[e]Significant label scrambling can be prevented with the use of DL39(DE3) or CT19 cells.
[f]Based on Marley et al. (2001). The cells are grown in LB, then resuspended in the labeled medium, and the protein expression is induced after 1 h incubation.

Wider and Wüthrich, 1999; Fiaux et al, 2002). The resulting approach enables the production of residue-specifically ^{15}N-labeled and perdeuterated proteins. The protocol has been applied to produce labeled samples of GroEL, the single-ring variant of GroEL, SR1, GroES, and the GroE substrate polypeptide human dihydrofolate reductase (hDHFR). The results are summarized and evaluated in the Tables 2 and 3.

The labeling protocol

The labeling scheme is based on a minimal medium (M9) supplemented with a limited amount of deuterated algal lysate amino acid mixture (Löhr et al., 2003; Markley et al., 1968; Markus et al., 1994; Serber et al., 2001; Yamazaki et al., 1997). Table 1 presents the medium composition, in which the ingredients marked with an arrow can be isotope-labeled in different combinations. The amount of protonated glucose contained in a liter of culture medium was reduced relative to the standard M9-medium to avoid endogenous biosynthesis of protonated amino acids; we thus obtained a higher degree of deuteration of the protein. The recipe is similar to commercial formulations for producing uniformly ^{15}N-labeled and deuterated proteins (for example, Table 2, line 3), and we produced uniformly labeled samples by using a [^{15}N,^{2}H]-labeled amino acid mixture and $^{15}NH_4Cl$ in our medium (Table 2, line 4). By limiting the amount of glucose and algal hydrolysate, this protocol can be adapted for the production of residue-specifically ^{15}N-labeled and deuterated proteins. For this purpose, deuterated algal lysate is used, and one or several [^{15}N,^{2}H]-amino acids are added in proper amounts

just before induction of protein expression. Although this is not further discussed here, ^{13}C-labeled components may also be used. In several variations, the proposed protocol meets the requirements for high overall protein deuteration, minimal scrambling of ^{15}N, and high protein yields (Tables 2 and 3, Figure 2).

Factors affecting the level of protein deuteration

Whereas partial deuteration may be useful for proteins in structure sizes of 30 to 100 kDa, complete replacement of non-exchangeable protons for deuterons is required to detect amide resonances in structures of molecular weight above 150 kDa. This applies also when ^{15}N-labeling is applied only to specified amino acid types in a protein. For example, in the first panel of Figure 1, we show that deuteration of only the ^{15}N-labeled residue is not sufficient to detect any resonances from structured parts of the 800 kDa chaperonin GroEL.

To achieve high deuteration, all solutions for the medium of Table 1 were prepared in D_2O, and the culture was inoculated from a preculture in the same fully deuterated medium. A deuteration level of $\sim$ 96% was achieved for samples of [^{15}N,^{2}H-Val, *U*-^{2}H]-GroES, [^{15}N,^{2}H-Lys, *U*-^{2}H]-GroES, and [^{15}N,^{2}H-Cys, *U*-^{2}H]-DHFR in spite of the use of 1 g/L of protonated glucose in the medium (Table 3, line 2). This deuteration level is comparable to that achieved with perdeuteration procedures which do not enable residue-specific ^{15}N-labeling (summarized in Table 2 for comparison). When the expressions were performed in H_2O-based medium, deuteration levels up to 90% were obtained for samples of GroES, GroEL or DHFR (Table 3, line 2) (see also Löhr et al., 2003, Markus et al., 1994, Serber et al., 2001, Yamazaki et al., 1997). This is of interest when growth of the host cells and protein expression are difficult in D_2O-based media. This variant of the protocol also ensures complete protonation of the amide groups, as compared with the expression in D_2O for which deuterons might be retained at buried amide sites, making them non-observable in NMR spectra based on observation of ^{1}H spins. Figure 1 shows that these schemes enable the observation of resonances originating from structured parts of the GroEL protein.

The deuteration level obtained when performing expression on H_2O-based medium correlates with the amount of deuterated amino acids present in the medium at the time of induction. We found that reducing the amount of algal lysate in the medium or delaying the induction resulted in lower overall deuteration of the expressed protein. As a general rule, the use of 3 g/L of algal lysate and induction in the mid-exponential phase ensures an overall deuteration level above 85%. This variant of the protocol uses a limited glucose concentration (1 g/L) in the medium, so that protonated glucose can be used to sustain the cellular energy metabolism without substantially affecting the final level of deuteration of the protein. Similarly, we observed that the extent to which α-deuterons in the amino acids are exchanged with solvent protons also depends on the amino acid concentrations in the medium (data not shown). This is consistent with earlier observations by others (Crespi et al., 1968; Löhr et al., 2003; Markus et al., 1994). For example, Löhr et al. (2003) showed that the fractional C^{α} deuteration can vary from 20 to 98% depending on the amino acid type. From high-resolution [^{15}N,^{1}H]-HSQC spectra of the protein samples denatured in 6 M guanidinium chloride or urea, we could assess that the C^{α} position of leucines in specifically ^{15}N-labeled samples were deuterated to more than 90% when following the aforementioned general recipe. Although the overall deuteration achieved with this protocol is comparable to that obtained from a cultivation on minimal medium in D_2O supplemented with protonated glucose, the distribution of the residual protons is expected to be different. In the latter case, glucose protons are incorporated mostly into the amino acid side-chains during biosynthesis (Neidhardt et al., 1996; Shekhtman et al., 2002), whereas in the H_2O-based cultivations described here the protonation will first affect the C^{α} position through the activity of transaminases. Clearly, the appropriate protocol for production of a uniformly labeled and partially deuterated sample should be selected according to the requirements of the NMR experiments planned (Löhr et al., 2003; Yamazaki et al., 1997).

An additional consideration for the production of highly deuterated proteins is the ability of a particular host strain and expression system to function under stringent conditions. Uniform ^{15}N-labeling and perdeuteration of proteins has been routinely performed by growing cells on ^{15}N-enriched, D_2O-based minimal medium supplemented with a deuterated carbon source. These conditions cause a significant stress to the bacterial cells, requiring long periods of adaptation of the cells to the growth medium. Some of our systems simply failed to grow in perdeuterated minimal medium. We found that cell growth and protein expression in a perdeuterated medium supple-

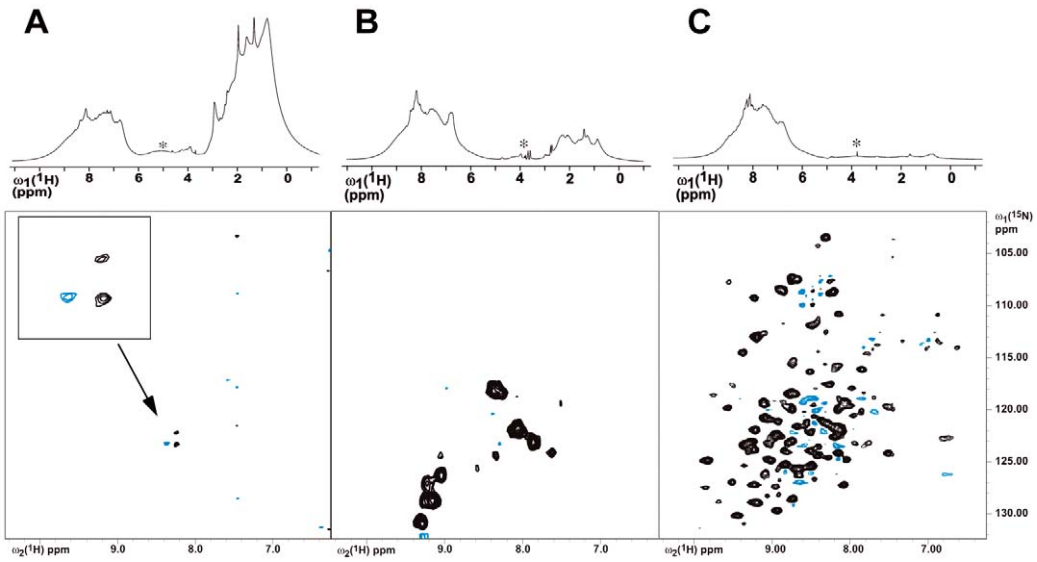

Figure 1. Effect of the level of deuteration on NMR spectra of uniformly and residue-specifically ^{15}N-labeled GroEL. A 1D ^{1}H-NMR spectrum (upper panel) and a [^{15}N,^{1}H]-CRIPT-TROSY spectrum (lower panel) are shown for three different preparations of GroEL. (A) [^{15}N,^{2}H-Leu]-GroEL prepared by addition of [^{15}N,^{2}H]-leucine to an unlabeled synthetic rich medium in H_2O (Table 3, line 1). Only the leucine residues are deuterated. The single resonance corresponds to residual intensity originating from Leu 531. Its multiplet structure is shown in more detail in the inset in the upper left corner. This residue is located in the flexibly disordered tail of the GroEL subunits, and its resonance is largely suppressed by the CRIPT transfer (Riek et al., 2002). The low deuteration level obtained using this labeling protocol is not sufficient to observe any resonances from the structured parts of GroEL. (B) [^{15}N,^{2}H-Leu, $\sim$90%^{2}H]-GroEL prepared by addition of [^{15}N,^{2}H]-leucine to a deuterated synthetic rich medium in H_2O (Table 3, line 2). This protocol enables observation of leucine resonances from structured parts of the large protein. Although there is some signal overlap, we estimate that the resonances of only 15 out of the total of 41 leucyl residues are detectable. Comparison with the spectrum (C) indicates that the remaining resonances are missing for reasons which are not related to the labeling scheme chosen (see also text). (C) [*U*-^{15}N, *U*-^{2}H]-GroEL prepared from minimal medium in D_2O containing $^{15}NH_4Cl$ as the sole nitrogen source and [$^{2}H_4$]-acetate as sole carbon source (Table 2, line 2). The overall deuteration level in this protein is $\sim$98%. The 1D ^{1}H-NMR spectra were recorded as described previously, with an acquisition time of 100 ms (Riek et al., 2002). The different CRIPT-TROSY spectra were recorded for 10 to 15 h, acquired data size 100*1024 complex points, $t_{1\,max}$ = 25 ms, $t_{2\,max}$ = 100 ms. The interscan delay was 0.3 s, and the CRIPT transfer delay was 1.4 ms (Riek et al., 2002). The samples were 1.5, 1.0 and 1.4 mM in GroEL subunits, respectively, in 25 mM potassium phosphate at pH 6.1, 20 mM KCl, 5% D_2O. The asterisks in the 1D ^{1}H-NMR spectra indicate artefactual peaks that do not belong to GroEL.

mented with ^{2}H-acetate often required the use of a very tightly regulated expression system, such as a T7 vector. However, the use of richer media, such as the one described here or commercially available isotope-labeled growth media, eliminated the requirement for pre-conditioning of the cells, which should ease the constraints relating to the expression system.

Factors affecting the residue-specific ^{15}N-labeling

Selective ^{15}N labeling of proteins by residue type typically involves the use of synthetic rich broth containing one or several ^{15}N-labeled amino acids, and high levels of all other amino acids in unlabeled form (McIntosh and Dahlquist, 1990; Muchmore et al., 1989). This approach suppresses the incorporation of the ^{15}N-label at undesired sites (cross-labeling) through metabolic pathways. Applying a similar procedure using deuterated ingredients to obtain a deuterated and specifically ^{15}N-labeled protein is currently hampered by limited commercial availability of the components. We therefore replaced the addition of individual labeled amino acids by adding a deuterated algal hydrolysate mixture, to which the ^{15}N,^{2}H-labeled amino acid is added. A similar approach has been described for selective protonation of a deuterated protein, resulting in simplified ^{1}H-NMR spectra (Markley et al., 1968). However, when applying this protocol to selective ^{15}N-labeling, cross-labeling becomes a substantial difficulty because the transfer of the amino group to other compounds is the first step in the amino acid metabolism. Careful adjustment of the amount of labeled amino acid is therefore crucial to prevent such cross-labeling. Since the deuterated algal hydrolysate already contains a pool of unlabeled amino acid, the quantity of ^{15}N-labeled amino acid

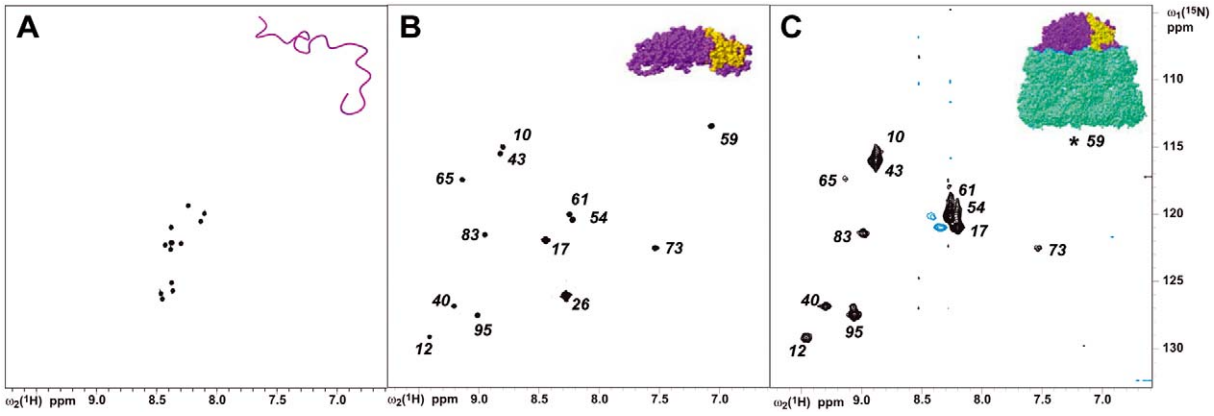

Figure 2. Residue-specific ^{15}N,^{2}H-labeling of deuterated GroES. (A) [^{15}N,^{1}H]-HSQC spectrum of denatured [^{15}N,^{2}H-Val, ∼96% ^{2}H]-GroES. The 13 resonances observed are those of the valine residues. No scrambling of the ^{15}N-label can be detected. The sample contained 0.6 mM GroES subunits, 6 M urea, 45 mM potassium phosphate at pH 6.1, 20 mM KCl, 8% D_2O, and protease inhibitors. The spectrum was recorded for 13 h, acquired data size 300*2048 complex points, $t_{1\,max}$ = 113 ms, $t_{2\,max}$ = 195 ms, the interscan delay was 1 s. (B) [^{15}N,^{1}H]-TROSY spectrum of folded [^{15}N,^{2}H-Val, ∼96% ^{2}H]-GroES. The numbers indicate the sequence-specific assignments obtained from TROSY-type triple resonance experiments with a ^{2}H,^{13}C,^{15}N-labeled sample (Fiaux, 2002). The sample contained 1.1 mM GroES subunits, 25 mM potassium phosphate at pH 6.1, 20 mM KCl, 8% D_2O, and protease inhibitors. The spectrum was recorded for 16 h, acquired data size 300*4096 complex points, $t_{1\,max}$ = 113 ms, $t_{2\,max}$ = 390 ms, the interscan delay was 1 s. (C) [^{15}N,^{1}H]-CRIPT-TROSY spectrum of [^{15}N,^{2}H-Val, U-^{2}H]-GroES in a 470 kDa complex with unlabeled SR1. 12 of the 13 valine resonances can be detected and were assigned by reference to the crystal structure, using NOESY data (unpublished). The asterisk indicates the position of the resonance of V59, which was observed in a [^{15}N,^{1}H]-CRINEPT-HMQC spectrum of the same sample. The spectrum was recorded for 5 h, acquired data size 100*1024 complex points, $t_{1\,max}$ = 25 ms, $t_{2\,max}$ = 81 ms. The interscan delay was 0.6 s, and the CRIPT transfer delay was 2.8 ms (Riek et al., 2002). In each panel an inset in the upper right corner represents a molecular model visualizing the conformational state of GroES. In the crystal structure of GroES (Hunt et al., 1996) in (B) one of the seven symmetry-related subunits has been colored in gold. The molecular model for the complex of GroES and SR1 in (C) is based on Xu et al. (1997).

to add must be sufficient to ensure reasonably high ^{15}N-occupancy in the protein while remaining small enough to avoid cross-labeling. Typically, 75 mg of ^{15}N-labeled leucine, 55 mg of ^{15}N-labeled valine, or 5 mg of ^{15}N-labeled cysteine were used per gram of algal lysate in the medium.

The importance of reducing cross-labeling cannot be over-emphasized when studying large structures with solution NMR. For example, even minor cross-labeling to other residues can lead to very strong signals in the NMR spectra when the receiving residues are located in flexible regions of the proteins. In an assignment procedure, it may then not be trivial to distinguish resonances arising from the labeled residue itself from those derived through cross-labeling. Multiply transaminase-deficient strains have been found to be suitable tools to suppress scrambling of the ^{15}N-label. For example, the *E. coli* strains DL39(DE3) (MG1655, F-, LAM-, aspC13, fnr-25, rph-1, ilvE12, tyrB507, λDE3), or CT19 (BL21, avtA::Tn5/, trpB83::Tn10/, aspC13, ilvE12, tyrB507, λDE3) eliminate cross-labeling via transaminase activity for valine, leucine, isoleucine, aspartate, phenylalanine, tyrosine and tryptophan residues (LeMaster and Richards, 1988; Waugh, 1996). Indeed, 'clean' ^{15}N-labeling of valine and leucine in our experiments was only achieved with these strains (data not shown). We expect that strains with genetic lesions in biosynthetic pathways shall also be useful for ^{15}N-labeling of amino acids that are direct metabolic precursors of other residues, for example, serine or aspartate (see LeMaster and Richards, 1988; Muchmore et al., 1989; Waugh, 1996 for detailed discussions of these aspects). For amino acids that lie at the end of biosynthetic pathways, such as arginine, cysteine, glycine, histidine, lysine, or methionine, labeling in an auxotrophic strain is not expected to be of significant advantage for preventing cross-labeling. In addition, dilution of the ^{15}N label by endogenous amino acid biosynthesis does not appear to be a serious difficulty. For example, in ascertaining the overall levels of deuteration, we found that endogenous amino acid synthesis contributes no more than a few percent of the amino acids when using our protocol. The adaptation of the protocol described by Marley et al. (2001) for the purpose of residue-specific labeling (see Table 3, line 4) may be an exception to these considerations: In this procedure, one can make use of the amino acid auxotrophies to promote

the turnover of non-labeled components before adding the labeled amino acids and inducing protein expression. Finally, note that residue-specific ^{15}N-labeling of glutamate or alanine will probably not be practicable due to the central metabolic role of these compounds.

Application to the study of the GroE proteins

We have used the procedures described here as well as other protocols from the literature to produce samples of the GroE proteins with different labeling patterns for solution NMR studies of structural and functional aspects of these chaperones. Tables 2 and 3 present an overview of the various samples used and the characteristic labeling patterns obtained.

In a first study, we prepared uniformly ^{15}N-labeled and perdeuterated GroEL (Table 2, line 2). The [^{15}N,^{1}H]-correlation spectra of this preparation (Figure 1C; see also Figure 3 in Riek et al., 2002) showed about 20% of the expected ^{15}N–^{1}H correlation peaks, which must include some resonances from structured parts of the protein. A sample was then produced with specific ^{15}N-labeling of the leucine residues, using a H_2O-based protocol (Table 3, line 2). This labeling pattern was designed to reduce spectral overlap and enable a more precise count of the resonances, as well as to investigate how many cross peaks are missing in the spectra due to slow $^{2}H^{N}/^{1}H^{N}$ back-exchange after protein expression in D_2O. In NMR studies of the resulting [^{15}N,^{2}H-Leu, $\sim$90%^{2}H]-GroEL, only about 15 out of the total of 41 leucine resonances could be detected (Figure 1B), suggesting that the remaining resonances are missing for reasons other than spectral overlap or incomplete $^{2}H^{N}/^{1}H^{N}$ exchange. For example, they might be broadened beyond detection by conformational averaging.

In contrast to GroEL, the complexes of ^{15}N-labeled and deuterated GroES with its partners GroEL or SR1 yielded virtually complete correlation spectra, and enabled a mapping of the interaction surface with these chaperones (Fiaux et al., 2002). We then produced residue-specifically ^{15}N-labeled and deuterated GroES as a tool for resonance assignment in the large complexes. Figure 2 shows correlation spectra of 70 kDa [^{15}N,^{2}H-Val, $\sim$96%^{2}H]-GroES free in solution (Figure 2B) and in the 470 kDa complex with SR1 (Figure 2C). The information obtained from residue-specific labeling combined with NOESY data, both for the free GroES and its complex with SR1, has enabled the assignment of most resonances in the complex (to be published).

Acknowledgements

This work was supported by the Schweizerischer Nationalfonds (projects 31.49047.96 and 31.49047.01), by HHMI and by NIH. We thank Krystyna Furtak for help in constructing the plasmids used for expression of GroEL, GroES and SR1. We thank Dr René Brunisholz, Fabrizia S. Sprecher and Dr Thomas Denziger for the mass spectrometry measurement, and Dr Hans Senn and Dr Beat Wipf at Hoffmann-LaRoche Inc., Basel, for a gift of the auxotrophic *E. coli* strain collection.

References

Bartels, C., Xia, T. H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1–10.

Crespi, H.L., Rosenberg, R.M. and Katz, J.J. (1968) *Science*, **161**, 795-796.

Fiaux, J. (2002) *Solution NMR with Large Macromolecular Assemblies: The GroE Chaperonin System*, PhD Thesis 14616, Eidgenössische Technische Hochschule Zürich, Zürich.

Fiaux, J., Bertelsen, E.B., Horwich, A.L. and Wüthrich, K. (2002) *Nature*, **418**, 207–211.

Gardner, K.H. and Kay, L.E. (1998) *Annu. Rev. Biophys. Biomol. Struct.*, **27**, 357–406.

Goldberg, M.S., Zhang, J., Sondek, S., Matthews, C.R., Fox, R.O. and Horwich, A.L. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 1080–1085.

Güntert, P., Dötsch, V., Wider, G. and Wüthrich, K. (1992) *J. Biomol. NMR*, **2**, 619–629.

Horwich, A.L., Burston, S.G., Rye, H.S., Weissman, J.S. and Fenton, W.A. (1998) *Meth. Enzymol.*, **290**, 141–146.

Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L. and Deisenhofer, J. (1996) *Nature*, **379**, 37–45.

Kelly, M.J., Krieger, C., Ball, L.J., Yu, Y., Richter, G., Schmieder, P., Bacher, A. and Oschkinat, H. (1999) *J. Biomol. NMR*, **14**, 79–83.

Leiting, B., Marsilio, F. and O'Connell, J. F. (1998) *Anal. Biochem.*, **265**, 351–355.

LeMaster, D.M. and Richards, F.M. (1988) *Biochemistry*, **27**, 142–150.

Lian, L.-Y. and Middleton, D.A. (2001) *Progr. Nucl. Magn. Reson.*, **39**, 171–190.

Löhr, F., Katsemi, V., Hartleib, J., Günther, U. and Rüterjans, H. (2003) *J. Biomol. NMR*, **25**, 291–311.

Markus, M.A., Dayie, K.T., Matsudaira, P. and Wagner, G. (1994) *J. Magn. Reson. B*, **105**, 192–195.

Markley, J.L., Putter, I. and Jardetzky, O. (1968) *Science*, **161**, 1249–1251.

Marley, J., Lu, M. and Bracken, C. (2001) *J. Biomol. NMR*, **20**, 71–75.

McIntosh, L.P. and Dahlquist, F.W. (1990) *Q. Rev. Biophys.*, **23**, 1–38.

Metzler, W.J., Wittelkind, M., Goldfarb, V., Mueller, L. and Farmer, B.T., II (1996) *J. Am. Chem. Soc.*, **118**, 6800–6801.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E. and Dahlquist, F.W. (1989) *Meth. Enzymol.*, **177**, 44–73.

Neidhardt, F.C., Curtiss, R., Ingraham, J.L., Liu, E.C.C., Low, K.B., Magasamik, B., Reznikoff, W.G., Riley, M., Schaechter, M.

and Umbarger, H.E. (Eds.) (1996) Escherichia coli *and* Salmonella typhimurium, 2nd edn., American Society for Microbiology, Washington, DC, USA.
Pervushin, K. (2000) *Q. Rev. Biophys.*, **33**, 161–197.
Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12366–12371.
Riek, R., Fiaux, J., Bertelsen, E.B., Horwich, A.L. and Wüthrich, K. (2002) *J. Am. Chem. Soc.*, **124**, 12144–12153.
Riek, R., Pervushin, K. and Wüthrich, K. (2000) *Trends Biochem. Sci.*, **25**, 462–468.
Riek, R., Wider, G., Pervushin, K. and Wüthrich, K. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 4918–4923.
Salzmann, M., Pervushin, K., Wider, G., Senn, H. and Wüthrich, K. (2000) *J. Am. Chem. Soc.*, **122**, 7543–7548.
Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning, a Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY.
Serber, Z., Ledwidge, R., Miller, S.M. and Dötsch, V. (2001) *J. Am. Chem. Soc.*, **123**, 8895–8901.
Shekhtman, A., Ghose, R., Goger, M. and Cowburn, D. (2002) *FEBS Lett.*, **524**, 177–182.
Smith, B.O., Ito, Y., Raine, A., Teichmann, S., Ben-Tovim, L., Nietlispach, D., Broadhurst, R.W., Terada, T., Kelly, M., Oschkinat, H., Shibata, T., Yokoyama, S. and Laue, E.D. (1996) *J. Biomol. NMR*, **8**, 360–368.
Venters, R.A., Farmer, B.T., II, Fierke, C.A. and Spicer, L.D. (1996) *J. Mol. Biol.*, **264**, 1101–1116.
Waugh, D.S. (1996) *J. Biomol. NMR*, **8**, 184–192.
Weissman, J.S., Hohl, C.M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A. and Horwich, A.L. (1995) *Cell*, **83**, 577–587.
Wider, G. and Wüthrich, K. (1999) *Curr. Opin. Struct. Biol.*, **9**, 594–601.
Xu, Z., Horwich, A.L. and Sigler, P.B. (1997) *Nature*, **388**, 741–750.
Yamazaki, T., Tochio, H., Furui, J., Aimoto, S. and Kyogoku, Y. (1997) *J. Am. Chem. Soc.*, **119**, 872–880.