In Vitro Production of Perdeuterated Proteins in H2O for Biomolecular NMR Studies

Lionel Imbert, Rachel Lenoir-Capello, Elodie Crublet, Alicia Vallet, Rida Awad, Isabel Ayala, Celine Juillan-Binard, Hubert Mayerhofer, Rime Kerfah, Pierre Gans, et al.

Running head Title:

Cell-Free Production of Perdeuterated Proteins

Title:

In Vitro Production of Perdeuterated Proteins in H_2O for Biomolecular NMR studies

Lionel Imbert[1,2,5], Rachel Lenoir-Capello[3,5], Elodie Crublet[4], Alicia Vallet[1], Rida Awad[1], Isabel Ayala[1], Celine Juillan-Binard[1,2], Hubert Mayerhofer[1], Rime Kerfah[4], Pierre Gans[1], Emeric Miclet[3], Jerome Boisbouvier[1,6]

1. Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71 avenue des martyrs, F-38044 Grenoble, France.
2. Univ. Grenoble Alpes, CNRS, CEA, EMBL, Integrated Structural Biology Grenoble (ISBG), 71avenue des martyrs, F-38044 Grenoble, France.
3. Sorbonne Université, École normale supérieure, PSL University, CNRS, Laboratoire des biomolécules, LBM, 75005 Paris, France
4. NMR-Bio, 5 place Robert Schuman, F-38025 Grenoble, France.
5. contributed equally
6. correspondence to be addressed to: jerome.boisbouvier@ibs.fr

Abstract:

The cell-free synthesis is an efficient strategy to produce in large scale protein samples for structural investigations. *In vitro* synthesis allows significant reduction of production time, simplification of purification steps and enables production of both soluble and membrane proteins. The cell-free reaction is an open system and can be performed in presence of many additives such as cofactors, inhibitors, redox systems, chaperones, detergents, lipids, nanodiscs and surfactants to allow the expression of toxic, membrane or intrinsically disordered proteins. In this chapter we present protocols to prepare *E. coli* S30 cellular extracts, T7 RNA polymerase and their use for *in vitro* protein expression. Optimizations of the protocol are presented for preparation of protein samples enriched in deuterium, a prerequisite for the study of high molecular weight proteins by NMR spectroscopy. An efficient production of perdeuterated proteins is achieved together with a full protonation of all the amide NMR probes, without suffering from residual protonation on aliphatic carbons. Application to the production of the 468 kDa TET2 protein assembly for NMR investigations is presented.

Key Words: Cell-free, *in vitro* protein synthesis, structural biology, isotopic labeling, NMR, perdeuteration.

1. Introduction

Nuclear magnetic resonance (NMR) spectroscopy is an established method to study interactions, structure, and dynamics of biomolecules at atomic resolution. This approach relies on the detection of NMR signals of natural hydrogen ^{1}H isotope, the most abundant nucleus in biomacromolecules, characterized by favorable NMR properties such as a high gyromagnetic ratio and a spin ½. Multidimensional ^{1}H-homonuclear NMR studies of proteins [1-2] were limited to systems with a molecular weight up to *c.a.* 10 kDa, due to the complexity of NMR spectra. During the last three decades, introduction of isotopic labeling techniques considerably increased the maximum size of biochemical systems that can be addressed by NMR spectroscopy. The use of robust protocols to introduce stable ^{15}N and ^{13}C spin ½ nuclei in recombinant proteins together with the development of triple resonance experiments have allowed spectroscopists to simplify and extend application of NMR to proteins with a molecular weight of *c.a.* 25 kDa [3-6]. Solution NMR studies of larger biomolecules are challenging due to the inherent spectral overlap between all the ^{1}H, ^{13}C and ^{15}N signals. Furthermore, rapid transverse relaxation (R_2) induces broadening and decreases intensity of NMR signals in high molecular weight proteins. This fast transverse relaxation is mainly due to the large number of intense dipolar interactions involving the abundant ^{1}H nuclei. As the magnitude of these dipolar interactions increases with the hydrodynamic radius of studied biomolecules, it is therefore more complex to study larger proteins characterized by slow overall tumbling.

Perdeuteration of proteins [7-10] was shown to improve relaxation properties in order to study larger targets. Deuterium isotope (^{2}H) has indeed a low gyromagnetic ratio, the dipolar interactions involving hydrogen nuclei are decreased by a factor of 43 when deuteron (^{2}H) substitutes the proton (^{1}H), leading to slower transverse relaxation and concomitant increase of sensitivity and resolution. In combination with optimized NMR experiments [11-12] using

spectrometers operating at high magnetic field, perdeuteration allows the study of monomeric proteins as large as 82 kDa [13] and, in favorable cases, complexes above 100 kDa [14-15]. Common methods to produce perdeuterated proteins for NMR studies usually rely on the overexpression of the target protein in *E. coli* grown in minimal media containing 100% $^{2}H_2O$ as solvent and a deuterated carbon source [10] (*see* **Figure 1A**). Such protocols enable protein perdeuteration up to 98 % [16]. In order to observe backbone amide protons ($^{1}H_N$), the overexpressed protein is purified and finally dialyzed against $^{1}H_2O$ to allow back exchange of protein $^{2}H_N$ with ^{1}H nuclei from the solvent. While this simple approach is suitable for fast exchanging protons, amide protons located in the core of large proteins exchange too slowly with the solvent to allow efficient back protonation. Generally, to reintroduce ^{1}H probes in these protected parts of the protein, the strategy consists in destabilizing the protein with chaotropic agents in $^{1}H_2O$ to speed up H_N exchange, before refolding the protein (*see* **Figure 1A**). The drawback of such a strategy is that the refolding of large proteins or membrane proteins is particularly challenging. At best, the target proteins will be refolded with poor yields and loss of precious labeled materials, but lot of proteins of biologic interest cannot be refolded *in vitro* in their native conformation in absence of cellular co-factors or chaperones, precluding production of a sample for NMR investigations.

The use of minimal medium prepared in 100% $^{1}H_2O$ buffer and supplemented with an excess of perdeuterated amino acids have been proposed to produce proteins fully labeled with ^{1}H nuclei on exchangeable sites of the overexpressed protein (*c.a.* 20 % of hydrogen in proteins), while non-exchangeable hydrogen covalently bound to aliphatic or aromatic carbons are deuterated at a level of c.a. 85% [17] (*see* **Figure 1B**). This level of deuteration allows acquisition of high quality 2D-(^{15}N, ^{2}H) spectra for large proteins [17-18]. The residual protonation is however not homogenously distributed on all the H_C, and a higher level of ^{1}H spin is observed on α sites (30% to 80% of ^{1}H [10]) because of abundant transaminases in living cells that are able to catalyze the

exchange of $^{2}H\alpha$ with protons from the solvent [10,19] (*see* **Figure 1B**). In small and medium size proteins, this heterogeneity results in the presence of extra $^{13}C_{\alpha}$ signals increasing complexity of NMR spectra, and in larger proteins the signals corresponding to $^{13}C_{\alpha}$–^{1}H isotopomers disappear due to the unfavorable transverse relaxation. Such artefacts introduced by the residual protonation deteriorate the quality of 3D NMR spectra used to sequentially assign backbone NMR frequencies, a prerequisite step for the investigation of proteins by NMR spectroscopy.

In vitro protein synthesis has been shown to be an attractive alternative method to produce perdeuterated proteins in $^{1}H_2O$ buffer (*see* **Figure 1C**). Cell-free, or *in vitro* expression of proteins, exploits transcription and translation machineries extracted from prokaryotic [20-21] or eukaryotic [22] cells. The cell extract, containing the required protein synthesis machineries, is recovered after removal of the DNA/RNA of the original organism. These machineries are supplemented with T7 RNA polymerase, amino acids and energy sources, in order to express the target protein from provided DNA template. Due to the absence of a biological membrane, the cell-free environment is an open system, offering the possibility of adding at any time compounds such as cofactors, ligands and stabilizers, in order to improve the synthesis of the target protein or to label the protein with isotopes without suffering from metabolic scrambling as observed *in vivo.* In this chapter, we describe protocols extensively applied by authors to produce *E. coli* S.30 extracts and T7 RNA polymerase, required for large scale *in vitro* synthesis of milligrams of protein samples for structural biology investigations. An optimized version of the protocol to produce perdeuterated proteins in $^{1}H_2O$ buffer for biomolecular NMR studies is also presented. Particular attention has been given to the quantification of residual protonation level on aliphatic carbons with and without treatment of S30 cell-free extract to inhibit transaminases activities. These protocols are illustrated by 2D-(^{1}H,^{15}N) solution and solid-state NMR spectra of peptidase TET2, a homododecameric protein assembly of 468 kDa. We show that *in vitro* expression of such large proteins in $^{1}H_2O$ supplemented with deuterated amino acids, allows recovery of a substantial number of signals for

important $^{1}H_N$ probes undetectable when proteins are produced using standard *in vivo* perdeuteration protocols.

2. Materials

The cell-free protein synthesis is a coupled reaction of transcription and translation starting from a DNA template. This step of transcription needs to be performed in RNAse free conditions. Careful consideration should therefore be given to the quality of plastic consumables and to the cleaning of glassware and devices used along the process. All the surfaces, pipettes and glassware have to be cleaned with RNAse remover (RNAse off, Shimitek), and washed with RNAse free water before being dried with ethanol. Users have to wear clean gloves all the time and work on ice as often as possible. All buffers should be prepared with RNAse free water, sterilized using a 0.22 µm filter.

2.1. *E. coli* extracts preparation (S30 extract)

1. Fermenter, protocol is described here for Techfors-S 42L (Infors HT)
2. RNAse free water (such as Direct Q3 with Biopak as final filter, Millipore)
3. Centrifuge J26 (Beckman) with JLA 9.1000, bottles (catalog ♯ 366751, Beckman)
4. Ultracentrifuge (Beckman) with 45Ti rotor and six 70 mL bottles (catalog ♯ 355655 Beckman)
5. Benchtop centrifuge for 1.5, 15 and 50 ml tubes
6. French press
7. OD_{600nm} spectrophotometer
8. Hybridization oven at 42°C or thermostatic orbital shaker
9. 4X Z-media: 165 mM KH_2PO_4, 664 mM K_2HPO_4, 40 g/L yeast extract
10. Glucose-thiamine solution: 240 g of glucose and 12 mg of thiamine per liter

11. *E. coli* strain BL21 DE3 or other.

12. 1 M Dithiothreitol **(**DTT)

13. 10X S30 buffer: 100 mM HEPES-KOH, pH 7.5, 600 mM $C_2H_3O_2K$; 140 mM $C_4H_6MgO_4$, add fresh DTT at 10 mM

14. Luria-Bertani (LB) medium for overnight culture

15. Antifoam 289 (Merck)

16. 5 M NaCl solution

17. Spectrapor 4 dialysis tube (12-14 MWCO) and magnetic-clamp

18. pIVEX GFP plasmid as model protein (RTS 100 *E. coli*, catalog ♯ BR1400106, Biotech rabbit)

2.2. T7 RNA polymerase expression and purification

1. 10 L of LB Ampicillin (100 mg/L final) for agar plates and culture medium.

2. 1 M Isopropyl-beta-D-thiogalactopyranoside (IPTG) solution (catalog ♯ EU0008-C, Euromedex)

3. *E. coli* strain BL21 DE3 (Invitrogen)

4. Vector containing the gene of the target T7RNApol (pAR1219)

5. Centrifuge J26 (Beckman) with 6 x 1 L rotor, JA 25.5 and JA 14 and dedicated bottles (catalog♯ 363678, 357002 and 355673, Beckman)

6. Ultracentrifuge Beckman with 45Ti rotor and six 70 mL bottles (catalog ♯ 355655, Beckman)

7. Ultrasonic Liquid Processors (Vibra cell VC505, SONICS)

8. FPLC device at 4°C (Duoflow 10, BioRad) with SP sepharose High Performance packed column (catalog ♯ 17108701, GE)

9. SDS-PAGE gels (Miniprotean gradient 4-20%, stainfree, Bio-Rad)

10. OD_{600nm} spectrophotometer

11. Lysozyme (catalog ♯ 5934C, Euromedex)

12. 0.8 % Sodium deoxycholate in water

13. 0.5 M NaOH and 2 M NaCl solution

14. 1 M DTT, 10 mg/mL Benzamidine, 0.1 M Phenylmethanesulfonyl fluoride (PMSF) in EtOH 100%

15. Buffer W: 20 mM Tris-HCl, pH 8.1, 20 mM NaCl, 2 mM EDTA, 1 mM DTT, 50 µM PMSF, 10 µg/mL Benzamidine

16. Spectrapor 4 dialysis tube (12-14 MWCO) and magnetic-clamp

17. Buffer R: 50 mM Tris-HCl, pH 8.1, 20 mM NaCl, 2 mM EDTA, 1 mM DTT, 50 µM PMSF and 10 µg/mL Benzamidine

18. 2 M Ammonium sulfate solution

19. 50 % Poly(ethyleneimine) (PEI) solution in H_2O (catalog ♯ 03880, Merck)

20. 10 % PEI solution in buffer A

21. 4.1 M Saturated ammonium sulfate solution, pH 7.0

22. 99 % Glycerol solution (catalog ♯ EU3550, Euromedex)

23. Buffer A: 20 mM NaH_2PO_4, pH 7.7, 50 mM NaCl, 1 mM EDTA, 1 mM DTT

24. Buffer B: 20 mM NaH_2PO_4, pH 7.7, 300 mM NaCl, 1 mM EDTA, 1 mM DTT

25. Buffer C: 20 mM NaH_2PO_4, pH 7.7, 100 mM NaCl, 1 mM EDTA, 1 mM DTT

2.3. Cell-free expression of proteins

1. 20 amino acids (Merck)
2. 1 M HCl (Merck)
3. 1 M KOH (Merck)
4. 100 mM solution of ribonucleotides CTP, GTP, UTP and ATP, pH 7.0 (catalog ♯ NU-1014, Jena Bioscience)
5. 2 M HEPES-KOH (Merck), pH 7.5

6. 10 mM folinic acid (Merck)

7. 100 mM 3',5'-cyclic AMP (Merck)

8. 1 M DTT (Merck)

9. 9.2 M ammonium acetate (Merck)

10. 1 M spermidine (Merck)

11. 1 M creatine phosphate (Merck)

12. 4 M potassium glutamate (Merck)

13. 1.07 M magnesium acetate

14. 17.5 mg/mL MRE600 tRNA (Roche)

15. 10 mg/mL creatine kinase (Roche)

16. Home-made T7RNApol (*see* **Section 3.2**)

17. Home-made S30 extract (*see* **Section 3.3**)

18. 1 µg/µL in water of target protein DNA cloned in pIVEX 2.3d or 2.4d vectors (RTS pIVEX *E. coli* vector set distributed by Biotechrabbit, catalog ♯ BR1400701)

19. Gebaflex dialysis device (Euromedex)

2.4. *In vitro* production of perdeuterated protein in H_2O

1. $NaBH_4$ (Merck)

2. Dimethylformamide (DMF)

3. 100 mM Acid Oxaloacetic (AOA), pH 7.5

4. 500 mM D-Malate (DM), pH 7.5

5. 10X S30 buffer: 100 mM HEPES-KOH, pH 7.5, 600 mM $C_2H_3O_2K$, 140 mM $C_4H_6MgO_4$, add fresh 10 mM DTT

6. Algal Amino Acids Mix (AAAM) as Celtone® (Cambridge Isotope Labeling) or Isogro® (Merck)

7. Protease from *Streptomyces griseus* (catalog # P5147, Merck)

8. 20 mM Tris-HCl, pH 7.5

9. Water-bath at 95°C

10. Benchtop centrifuge for 1.5, 15 and 50 mL tubes (Eppendorf)

3. Methods

3.1. *E. coli* extracts preparation (S30 extract)

S30 extracts can be prepared in shake-flasks or a fermenter. (*see* **Note 1**). The protocol detailed herein corresponds to a 12 L-fermenter culture of *E. coli* BL21 DE3 strain, enabling the preparation of 200 mL of S30 extract (*see* **Note 2**). RNAses contaminations can occur anytime during the protocol, so strict cleaning procedures with RNAse remover should be applied for glassware and RNAse free plastic consumables should be used. Moreover, as ribosomes are temperature sensitive, S30 extract preparation must be performed on ice with chilled glasswares/plastics.

1. Fermenter preparation: The day before the culture, prepare all the solutions listed in the dedicated Materials section, autoclave the glucose-thiamine solution and a funnel. Fill the fermenter tank with 9 L of H_2O and 3 L of 4X Z-media and autoclave the 12 L *in situ* (*see* **Note 3**). Inoculate 400 mL of LB with 50 µL of BL21 DE3 cells (glycerol stock) at 37°C overnight.
2. *E. coli* culture: In the morning, warm-up the medium at 37°C, set up the stir at 550 rpm and the airflow at maximum level (*see* **Note 4**). Add the 1 L glucose-thiamine solution with a sterile funnel and inoculate the fermenter with the 400 mL overnight culture at an initial OD_{600} of 0.1. Follow the pH, pO_2 and OD_{600} every 30 minutes. A decrease of pO_2 is

expected together with an increase of the optical density. Culture should be stopped when the value of pO_2 is close to zero or the OD_{600} equal to 3, usually occurring 3 to 4 hours after the beginning of the culture (*see* **Note 5**). The temperature must be quickly reduced from 37 to 16°C, and the OD_{600} should not exceed 3.2 (*see* **Note 6**). Harvest the cells by centrifugation for 15 minutes at 5,000 *g* and 4°C.

3. Preparation of S30 extracts.

Wash: Perform 3 washing steps of the cell pellets with cooled S30 buffer, the first one with 2 liters, the second one with 1 liter and finally with 0.5 Liter. At each step, pellet the cells at 4°C for 15 minutes at 5000 *g* (*see* **Note 7**). Weigh the wet cells and store the pellet overnight on ice in a cold room (*see* **Note 8**).

Lysis: The next day, cells are resuspended in cold S30 buffer (1.27 mL of buffer per gram of wet cells) and then disrupted using French press (only one pressure cycle). The supernatant is clarified by two centrifugation steps at 4°C in six 70 ml ultracentrifuge tubes, using a 45 Ti rotor (30,000 *g*, 30 minutes). After each centrifugation step, recover only the supernatant (corresponding to 80 % of the initial volume).

Maturation: Process the endogenous mRNA by incubating the S30 extract supernatant at 42°C for 45 minutes after the addition of a 5 M NaCl solution to reach a final concentration of 400 mM [23].

Dialysis: Glassware (a 2 L cylinder as dialysis tank, a 500 mL one for measuring the S30 volume), S30 buffer (12 L prepared without DTT) and a 12 kDa cut-off dialysis membrane should be stored in advance at 4°C for dialysis (*see* **Note 9**). S30 extract is dialyzed in several steps

each time against 2 L of S30 buffer at 4°C. The DTT (1 mM) should be added at the beginning of each dialysis step. Perform two successive 1h-dialysis, one overnight dialysis bath and three extra 1h-dialysis steps the next morning. Centrifuge the S30 extract in 50 ml tubes at 5000 *g* and 4°C. Prepare 2 mL safe-lock tubes to aliquot the supernatant in 1 mL fractions and freeze the tubes immediately in liquid nitrogen. These S30 extract aliquots can be stored for several years at -80°C.

Quality control: Expression yields are very dependent on Mg^{2+} concentration and the precise Mg^{2+} concentration already present in the S30 extract can vary from batch to batch. For each S30 extract batch, the Mg^{2+} concentration has to be optimized by performing expression tests, in duplicate, using a model protein. We usually perform six *in vitro* expression tests in small volumes, with additional magnesium concentration in the reaction and feeding mixes ranging from 5 to 15 mM. This optimization is routinely performed using GFP as a model protein, and we determine the optimal concentration of Mg^{2+} from highest fluorescence intensity.

3.2. T7 RNA polymerase expression and purification

T7 RNA polymerase (T7RNApol) needs to be purified in RNAse free conditions at 4°C. It is an essential component of the Cell-free reaction and high concentrations are needed for the reaction. It can be purchased from different providers or produced in-house by overexpression in *E. coli* and a one-step ion exchange purification (*see* **Note 10**) [24-25].

1. Culture: Spread BL21DE3 cells transformed with the plasmid pAR1219 coding for T7RNApol on a LB Ampicillin (LBA) agar plate and incubate overnight at 37°C. Inoculate a 30 mL LBA culture with a single colony and let grow for 8 hours at 37°C with vigorous

shaking. In the evening, inoculate an overnight culture of 200 mL at OD_{600}: 0.1. The next day, warm up 10 L at 37°C of culture medium for overexpression of T7RNApol (ten 3L flasks filled with 1 L of LBA media). Inoculate with the overnight culture at OD_{600}: 0.1, let grow and induce with IPTG at 0.3 mM when OD_{600} reaches 0.8. After 2 hours at 37°C, collect the cells by centrifugation (5,000 *g*, at 4°C for 15 minutes). In order to centrifuge only 5 L of culture simultaneously, the inoculation for 5 out of the 10 flasks can be shifted from 30 minutes. Wash the cells with buffer W and pellet the cells by centrifugation. Weigh the wet cells, and store pellets on ice overnight in a cold room. Around 40 g of wet cells are expected from a 10 L culture using this protocol.

2. Lysis: resuspend the pellet with 2.88 mL of buffer R per gram of wet cells, for instance 115.2 mL for a 40 g cell pellet. The lysis is initiated by adding 1.5 mg of lysozyme per gram of wet cells, followed by an incubation of 20 min on ice with occasional shaking. The detergent sodium deoxycholate (0.8 % stock solution in water) is subsequently added at 0.3 mL/g of wet cells (*i.e.* 12 mL for a 40 g pellet) and incubated for 20 min on ice with occasional shaking. The viscosity is reduced by sonication in a beaker on ice. The supernatant is clarified by two centrifugation steps at 4°C. The first step is performed at 10,000 *g* for 15 min, and the supernatant is then ultracentrifugated for 3 hours at 140,000 *g*.

3. Ammonium sulfate precipitation: Adjust the supernatant volume to 5.76 mL per gram of wet cells with buffer A in a chilled beaker. Addition of a 2 M ammonium sulfate solution is performed slowly, drop by drop, on ice under stirring up to a final concentration of 0.2 M. To initiate precipitation of high molecular weight nucleic acids, add slowly, with stirring, on ice, PolyEthyleneImine at a final concentration of 0.5%. Incubate for 20 min on

ice in a rocker-shaker. Centrifuge for 10 min at 30,000 *g*, at 4°C, collect the supernatant and measure its volume (noted V) in a chilled cylinder (*see* **Note 11**). The T7RNApol is precipitated using a drop-by-drop addition of an ammonium sulfate saturated solution (total volume added 0.82xV). The addition is performed in a centrifuge bottle under stirring on ice, and followed by extra stirring for 15 minutes. The sample is centrifuged for 20 min at 15,000 *g* and 4°C, and the pellet is solubilized with 4 mL of buffer C per gram of wet cells. In order to eliminate ammonium sulfate, the solution is dialyzed once against 5 L of chilled buffer A, using a 12 kDa cut-off membrane.

4. Ion exchange purification (*see* **Note 12**): the supernatant is clarified by centrifugation for 10 minutes at 30,000 *g* and loaded at a flow rate of 2-3 mL/min on the SP-sepharose column, previously equilibrated with buffer A. The column is washed with 1.5 column volume (CV) of buffer A and the T7RNApol is eluted by a gradient from 0 to 100% of buffer B in 20 CV. The enzyme is recovered around 100 mM NaCl and collected by 5 mL fractions. The fractions are analyzed on SDS-PAGE gel and the ones containing T7 RNApol at highest concentration (center of the chromatography peak) are pooled, and dialyzed against 1 L of cold buffer C containing 50% glycerol. After an overnight dialysis, the volume naturally decreases by 2 (because of the glycerol), to reach a concentration of *c.a.* 5 mg/mL. The T7RNApol can be stored for at least one year at -20°C (*see* **Note 13**).

3.3. Cell-free expression of proteins

We present here the Continuous Exchange Cell-free (CECF) reaction using a dialysis membrane system [26-27]. This protocol is optimized for the large-scale expression of unlabeled proteins typically used in structural biology projects. This section describes the preparation of the reaction

mix (1 mL) and the feeding mix (10 mL) required to synthesize *in vitro* the target protein. Particular attention should be paid to the design and purity of the DNA vectors coding for the target protein. pIVEX vectors [28] have been developed by Roche and are optimized for *in vitro* expression using T7RNApol and *E. coli* S30 extracts. They are designed with a T7 promoter, Ribosome Binding Site, T7-terminator sequences to allow efficient protein synthesis, His-tag sequence at N or C-terminus to facilitate detection or purification of expressed protein, and Multi Cloning Sites for insertion of target protein DNA sequence. The DNA sequence for N-terminal His-tag is optimized to ensure efficient initiation of transcription and translation. Alternative optimized expression tags [29] can be incorporated to increase production yield of the target protein. The DNA vector should be prepared in large scale (*c.a.* 500 µg) using a commercial plasmid preparation kit and eluted in RNAse free water at a concentration of 1 mg/mL. In the following protocol, vortex all the mixtures, except at point 6.

1. Solubilize amino acid mixtures in three 50 mL tubes, the concentration of each amino acid being 50 mM. Alanine, Arginine, Glycine, Histidine, Lysine, Proline, Serine, Threonine, and Valine are resuspended together in water. Acidic soluble amino acids (Asparagine, Aspartate, Cysteine, Glutamine, Glutamate, Leucine, Methionine, Tryptophan and Tyrosine) are resuspended in HCl 1 M. Isoleucine and Phenylalanine are resuspended in KOH 1 M.
2. Wash extensively the dialysis Gebaflex tube with water and keep it in water during the CECF reaction preparation (*see* **Note 14**).
3. Prepare the amino acid solution (AA mix) by adding 247.5 µL of each amino acid mix (water-soluble, acid-soluble and base-soluble amino acids), complement with 82.5 µL RNase free water.

4. Prepare the 10X reaction mix containing HEPES/KOH (55 mM, pH 7.5), DTT (3.4 mM), ATP (1.2 mM), 3', 5'-cyclic AMP (0.64 mM), 0.8 mM of each CTP, GTP and UTP ribonucleotides, folinic acid (68 μM), ammonium acetate (27.5 mM) and spermidine (2 mM).
5. Prepare the feeding mix with 1.05 mL of the 10X reaction mix, creatine phosphate at final concentration of 80 mM, Potassium Glutamate (208 mM), AA mix (1 mM for each amino acid), and magnesium acetate (14.4 mM).
6. Prepare the reaction mix as described for the feeding mix, with further addition of 250 mg/mL of creatine kinase, 175 μg/mL of total tRNA E. coli MRE600, 50 μg/mL of T7 RNApol, 40% of the final volume of S30 extract and 16 μg/mL of the target protein vector. Do not vortex the reaction at this stage.
7. Load the reaction mix (1 mL) in the dialysis Gebaflex and the feeding mix (10 mL) in a 25 mL cylinder and incubate overnight with stirring at a temperature optimal for the target protein (*see* **Note 15**).
8. Next morning, recover the supernatant by centrifugating 20 minutes at 10,000 *g* and proceed to purification according to the standard protocol of the protein. A dilution by a factor 4 is usually required in order to decrease sample ionic strength before loading it on a purification column.

3.4. *In vitro* production of perdeuterated protein in H_2O

The above described protocol enables mgs scale protein production for structural investigations. Dynamics and structural studies of proteins using NMR require uniform enrichment of the protein with stable ^{15}N and ^{13}C isotopes. Such uniform labeling schemes are easily obtained using protocols described in part 3.3 by substituting unlabeled amino acids (Table 1) by U-(^{15}N), or U-(^{15}N,^{13}C)-labeled mix of amino acids or labeled algal extract available from isotope suppliers (*see* **Note 16**

and **Note 17**). With the protocol described in this chapter, we typically obtain U-(^{15}N), or U-(^{15}N,^{13}C)-labeled protein in CECF mode with a yield of *c.a.* 2 mg/mL of reaction mix. Production of perdeuterated proteins can be achieved by substitution of $^{1}H_2O$ by $^{2}H_2O$ solvent [30], lyophilization of all components used in the cell-free reaction (Table 1 with exception of T7RNApol) and use of perdeuterated amino acids mix (*see* **Note 16**). Experimentally, we observed that performing cell-free reactions in $^{2}H_2O$ buffer reduces the protein yield by a factor of *c.a.* 2. Furthermore, for NMR studies of large proteins, the destabilization of proteins with chaotropic agents is required to allow back-exchange of the backbone $^{1}H_N$ proton. Prior to structural investigation, the protein needs to be refolded in its native fold and this tedious step is usually associated with a low protein recovery yield. Alternatively, the use of U-(^{2}H,^{15}N) or U-(^{2}H, ^{15}N,^{13}C)-labeled amino acids for *in vitro* protein synthesis in $^{1}H_2O$ solvent results in perdeuterated proteins fully protonated on all the exchangeable hydrogen positions [30] without requiring protein refolding. However, in this case, as the S30 extract contains active transaminases, perdeuterated amino acids added for the cell-free reaction are processed by enzymes catalyzing the exchange of α-deuteron by the solvent proton [30-32]. Although the level of residual protonation is reduced compared with the corresponding one observed *in vivo* using $^{1}H_2O$ and perdeuterated amino acids (30 to 80 % of residual protonation is observed on C_α [10,19]), it remains substantial at an average level of *c.a.* 25 % with values varying for the different amino acids from a few percent up to *c.a.* 50 % (*see* **Figure 2**). Such residual protonation on aliphatic protons on backbone atoms introduces heterogeneities in the samples and deteriorates the quality of most NMR spectra. To limit such residual protonation, the transaminases can be inactivated by the addition of inhibitors (AOA, D-malate) and $NaBH_4$ reduction of PLP co-factors with the following protocol [31-32]. During its preparation, the S30 extract can be treated with $NaBH_4$ after the first dialysis step (*see* **Section 3.1.3**) and then stored similarly as the untreated extract. Residual protonation on Cα site in proteins produced *in vitro* using this treated S30 extract is reduced, with a maximum residual protonation

of 10 % for few amino acids, and an overall mean of less than 5% residual protonation (*see* **Figure 2**). The $NaBH_4$ treatment and the addition of AOA and D-malate in the S30 are compatible with the cell-free protein production but usually result in a decrease of protein synthesis yield by a factor *c.a.* 2.

Inhibition of transaminases from S30 extracts

1. Solubilize the $NaBH_4$ powder at 100 mM in dimethylformamide (DMF) to keep its reductive power (*see* **Note 18**).
2. Treat the S30 extract by addition of $NaBH_4$ (at a final concentration of 20 mM) in a large container under gentle shaking at 4°C and incubate the reaction for 10 minutes.
3. $NaBH_4$ and DMF are removed from the S30 extract by three dialysis against 100 volumes of cold S30 buffer using 12 kDa cut-off dialysis tubes.
4. Centrifuge the S30 extract in 50 mL tubes at 5000 *g* and 4°C. Prepare 2 mL safe-lock tubes to aliquot the supernatant in 1 mL fractions and freeze the tubes immediately in liquid nitrogen. These $NaBH_4$ treated S30 extract aliquots can be stored for several years at -80°C.

In vitro synthesis of perdeuterated protein

In order to produce perdeuterated proteins with minimal residual protonation on aliphatic hydrogen sites in CECF mode, the following modifications need to be implemented in the protocol described in section 3.3. The three unlabeled amino acids mixtures (step 1) should be replaced by an adequate mix of U-(^{15}N,^{2}H) or U-(^{15}N,^{13}C,^{2}H) amino acids (*see* **Note 16**). The total volume of amino acids should be less than 825 µL (step 3), and the average concentration of each amino acid should be 15 mM (or 1 mM in final reaction or feeding mix). In order to avoid incorporation of unlabeled glutamate, the 208 mM of potassium glutamate (step 5) should be substituted by 150 mM of

ammonium acetate. $NaBH_4$ treated S30 extract should be used at step 6, and AOA and DM should be added in S30 extracts at a concentration of 20 mM each following the protocol described above.

3.5. Application to the production of large perdeuterated proteins for NMR investigations

The introduction of advanced isotopic labeling protocols, based on perdeuteration and selective protonation of few sites in the target protein, has enabled NMR investigations of very large protein assemblies [15, 33-35]. Recently, the near complete assignment of backbone heavy atoms of the 468 kDa homododecameric protein TET2 [36-37] was obtained by solid state NMR [38-39]. To accomplish this *tour de force*, Schanda and coll. had to use several uniformly and specifically labeled samples, including perdeuterated samples. Surprisingly, despite that most frequencies of the backbone heavy atoms were assigned, less than half of the backbone ^1H_N nuclei frequencies could be identified. One of the factors limiting the assignment of amide proton frequencies is the high stability of this protein. Indeed, as the used samples were produced in M9/^2H_2O *E. coli* cultures (*see* **Figure 1A**), all hydrogen positions are thus deuterated and many of those located in the core of such stable protein are protected from the exchange with the buffer solvent, hampering back-protonation when TET2 is dialyzed against ^1H_2O. The complete TET2 denaturation followed by its refolding in ^1H_2O solvent could represent a solution [34]. However, TET2 refolding is a tedious process, characterized by a very low protein recovery yield, precluding cost-effective production of such sample.

In order to protonate all the amide sites, we decided to produce perdeuterated TET2 protein directly in a ^1H_2O buffer using the cell-free synthesis protocol described above. The corresponding 2D-(^1H, ^{15}N) solid state NMR spectrum (*see* **Figure 3**) displays well dispersed signals characteristic of folded proteins. In order to investigate the effects of incomplete back-protonation, a second sample has been produced using the same protocol but using ^2H_2O as a solvent during *in vitro* protein synthesis. The figure 4A-B presents the same zoom of a 2D-(^1H, ^{15}N)-CRINEPT-

HMQC-TROSY [12] acquired in solution using both samples. A lot of amide proton signals observed when the sample is produced in $^{1}H_2O$ solvent (*see* **Figure 4A**) cannot be observed anymore when the sample is produced in $^{2}H_2O$ (*see* **Figure 4B**), although the sample was previously incubated in $^{1}H_2O$ for 10 days to promote exchange of amide deuterons with solvent protons. Similarly, the sample used to acquire spectrum displayed on figure 4A was dialyzed in $^{2}H_2O$ for 10 days, before acquiring a new 2D -(^{1}H, ^{15}N) spectrum (*see* **Figure 4C**). Most of the missing signals for the sample produced in $^{2}H_2O$ (*see* **Figure 4B**) can still be observed with the sample produced in $^{1}H_2O$ but extensively dialyzed in $^{2}H_2O$ (*see* **Figure 4C**). These experiments confirm that incomplete back-protonation of amide protons in large perdeuterated proteins is a major drawback for their investigations using NMR spectroscopy. As shown in this chapter, cell-free protein synthesis offers an efficient alternative to produce fully perdeuterated proteins with all solvent-exchangeable hydrogen sites fully occupied with ^{1}H spins, and without suffering from protonation artifacts on other sites. S30 extracts can be prepared in large scale, aliquoted and stored at -80 °C, before being used to produce in 24 h perdeuterated proteins in milligram quantities. After purification, such samples can be used directly to collect NMR data, without requiring tedious refolding of the perdeuterated proteins in $^{1}H_2O$ to ensure protonation of all NMR observable amide sites.

4. Notes

1. Alternatively, the 12 L culture could be performed in flasks up to an OD_{600}= 1, enabling the preparation of 50 mL of S30 extract.
2. Different *E. coli* strains can be used for S30 extract preparation, such as A19 [27], BL21 DE3 star [40], Rosetta pRARE [41], codon-plus RIL [42].

3. In this protocol, we use a Techfors-S fermenter (Infors HT) equipped with a steam generator allowing *in situ* sterilization. Note that approximately 1 L of water is lost by evaporation during the autoclave of the culture medium. If the fermenter is not equipped with a steam generator, use instead sterile water (8 L) and 3 L of 4X Z-media autoclaved separately.
4. Wait for 20 minutes before calibrating the 100% pO_2 at 37°C, 550 rpm and at maximum airflow.
5. If pO_2 starts to increase, bacteria will not be in the exponential phase anymore and ribosomes will have a lower activity. For BL21(DE3), final OD_{600} should not exceed 3.2 (and culture should not last for more than 4 hours to stay in the exponential growth phase). These values must be adapted for the different *E. coli* strains.
6. The complete cooling needs to be achieved in less than 20 minutes. Use chilled water flow at a temperature of less than 16°C to cool down the double-walled fermentation vessels.
7. The pellet can easily be resuspended using a cell homogenizer such as Tissue Master 125 (OMNI International).
8. The total mass of wet cells for a 12 L culture is expected to be around 150 g. We recommend a short storage of cell pellets at 4°C rather than freezing the cells.
9. Prepare the dialysis buffer in advance without DTT and store it at 4°C. Add the DTT extra temporally.
10. If the T7 RNApol has a His-tag, it will co-elute with the His-Tagged target protein during affinity purification. It has to be taken into account when the purification strategy is defined.
11. Precipitation of nucleic acids can be monitored by measurement of absorbance at 260 nm before and after addition of PEI.

12. The chromatography device needs to be RNAse free. Flush extensively all the system with 0.5 to 1 M NaOH, RNAse free water, 2 M NaCl solution and finally RNAse free water.
13. Analysis of T7RNApol (MW: 98800 Da, molar ε: 140,000 $M^{-1}.cm^{-1}$) is usually performed on 8% SDS-PAGE gel and by absorbance measurement.
14. Cut-off should be 12 kDa or less depending on the size of synthesized protein.
15. If protein is degraded or precipitated during cell-free expression, optimization of the duration and temperature of the synthesis should be performed to determine the optimal conditions.
16. The sources of labeled amino acids (AA) with different labeling schemes can be found from many providers (CIL, Merck, Cortecnet, …). They are available as individual amino acids, 16 AA mix, or 20 AA mix ready-to-use for cell-free expression. The main advantage of individual amino acids is that users can adjust the concentration of each amino acid in accordance with the protein target sequence [43]. Algal extract (Isogro® or Celtone®) can also be used as a cheap source of 16 labeled amino acids. C, W, N and E amino acids are degraded during Algal extract preparation and need to be added in the amino acid mix for cell free expression. Algal extracts contain *c.a.* 60% of these 16 AA at different concentrations, in the form of free amino acids or small peptides. These small peptides can be enzymatically hydrolyzed by a protease (pronase from *Streptomyces griseus*) in order to be incorporated during *in vitro* protein synthesis. To hydrolyze 1 g of algal extract in a volume of 20 mL (20 mM Tris-HCl pH 7.5), add 33 mg of this protease and incubate under stirring at 37°C overnight. The next morning, proceed to a heat shock (95°C for 30 minutes) to inactivate the protease. Centrifuge (5000 g, 15 minutes) and recover the supernatant containing the free amino acids. Hydrolyzed algal extract should be used at a final concentration of 3-5 mg/mL in the feeding and reaction mixtures.

17. In order to reduce the amount of expensive labeled amino acids, the cell-free reaction can be performed in Batch mode [27]. Synthesis in Batch mode is performed using only the reaction mix described in section 3.3. Cell-free reaction by-products inhibit the protein synthesis and therefore the reaction is usually limited to 2-3 hours. This protocol reduces quantities of labeled amino acids required by a factor of 10 but also decreases the protein synthesis yield by a factor of 3-5.
18. As $NaBH_4$ is a strong reducing agent activated by water, the stock solution must be prepared in DMF. $NaBH_4$ attacks the aldehyde group of PLP and reduces the Schiff base occurring between PLP and NH group of transaminases. H_2 release is accompanied by the formation of foam. Work under a fume-hood.

Acknowledgments: The authors thank Prof. Eva Pebay-Peyroula for providing the clone of the HMA8 ATPases nucleotide-binding domain from *Arabidopsis thaliana*. This work used the high field NMR and Cell-Free facilities at the Grenoble Instruct-ERIC Center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work was supported by grants from CEA/NMR-Bio (research program C24990) and the Agence Nationale de la Recherche (ANR-17-CE29-0010 CH2-PROBE).

5. References

1. Wüthrich K (1986) NMR of Proteins and Nucleic Acids. Wiley, New York
2. Bax A (1989) Two-dimensional NMR and protein structure. Annu Rev Biochem 58:223-56.

3. Ikura M, Kay LE, Bax A (1990) A novel approach for sequential assignment of 1H, 13C, and 15N spectra of larger proteins: Heteronuclear triple-resonance NMR spectroscopy. Application to calmodulin. Biochemistry 29:4659-4667
4. Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple resonance NMR spectroscopy of isotopically enriched proteins. J Magn Reson 213: 442-445
5. Bax A (2011) Triple resonance three-dimensional protein NMR: Before it became a black box. J Magn Reson 89:496-514
6. Ikura M, Clore GM, Gronenborn AM, Zhu G, Klee CB, Bax A (1992) Solution structure of a calmodulin-target peptide complex by multi-dimensional NMR. Science 256:632-638
7. Yamazaki T, Lee W, Arrowsmith CH, Muhandiram DR, Kay LE (1994) A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of 15N, 13C, 2H Labeled Proteins with High Sensitivity. J Am Chem Soc 116: 11655-11666 (Article)
8. Venters RA, Huang C-C, Farmer BT II, Trolard R, Spicer LD, Fierke CA (1995) High-level 2H/13C/15N labeling of proteins for NMR studies. J Biomol NMR 5:339-244
9. Muchmore SW, Sattler M, Liang H, Meadows RP, Harlan JE, Yoon HS, Nettesheim D, Chang BS, Thompson CB, Wong SL, Ng SL, Fesik SW (1996) X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. Nature 381:335-341
10. Gardner KH, Kay LE (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct. 27:357-406.
11. Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94:12366-71
12. Riek R, Wider G, Pervushin K, Wüthrich K (1999). Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. Proc Natl Acad Sci U S A. 96:4918-23

13. Tugarinov V, Muhandiram R, Ayed A, Kay LE (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase g. J Am Chem Soc 124:10025-35
14. Wider G, Wüthrich K (1999) NMR spectroscopy of large molecules and multimolecular assemblies in solution. Curr Opin Struct Biol 9:594-601
15. Fiaux J, Bertelsen EB, Horwich AL, Wüthrich K (2002) NMR analysis of a 900K GroEL GroES complex. Nature. 418:207-211
16. Sounier R, Blanchard L, Wu Z, Boisbouvier J (2007) High-Accuracy Distance Measurement Between Remote Methyls in Specifically Protonated Proteins. J Am Chem Soc 129: 472-473
17. Fiaux J, Bertelsen EB, Horwich AL, Wüthrich K (2004) Uniform and residue-specific 15N-labeling of proteins on a highly deuterated background. J Biomol NMR 29:289-297
18. Rasia R, Noirclerc-Savoye M, Gallet B, Bologna N, Plevin M, Blanchard L, Palatnik J, Brutscher B, Vernet T, Boisbouvier J (2009) Parallel screening and optimization of protein constructs for structural studies. Protein Science 18:434-439 (2009)
19. Smith BO, Ito Y, Raine A, Teichmann S, Ben-Tovim L, et al. (1996) An approach to structure determination using limited NMR data from larger proteins selectively protonated at specific residue types. J Biomol NMR 8:360–368
20. Zubay G (1973) In vitro synthesis of protein in microbial systems. Annual Rev. Genet. 7:267-87
21. Pratt C (1980) Kinetics and regulation of cell-free alkaline phosphatase synthesis. Journal of Bacteriology 143(3):1265-74
22. Roberts BE, Paterson BM (1973) Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. PNAS 70:2330–2334

23. Klammt C, Löhr F, Schafer B, Haase W, Dötsch V, Ruterjans H, Glaubitz C, Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. Eur J Biochem 271:568–580. doi: 10.1111/j.1432-1033.2003.03959.x
24. Davanloo P, Rosenberg AH, Dunn JJ, Studier FW (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase. Proc Natl Acad Sci U S A 81:2035-2039
25. Zawadzki V, Gross HJ (1991) Rapid and simple purification of T7 RNA polymerase. Nucleic Acids Res. 25:1948
26. Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. Science 242:1162-1164
27. Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T, Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. FEBS Lett 442:15-19
28. Martin GA, Kawaguchi R, Lam Y, DeGiovanni A, Fukushima M, Mutter W (2001) High-Yield, In Vitro Protein Expression Using a Continuous-Exchange, Coupled Transcription/Translation System. BioTechniques 31:948-953
29. Haberstock S, Roos C, Hoevels Y, Dötsch V, Schnapp G, Pautsch A, Bernhard F (2012) A systematic approach to increase the efficiency of membrane protein production in cell-free expression systems. Protein Expression and Purification 82:308-316
30. Etezady-Esfarjani T, Hiller S, Villalba C, Wüthrich K (2007) Cell-free protein synthesis of perdeuterated proteins for NMR studies. J Biomol NMR. 39:229-238
31. Su XC, Loh CT, Qi R, Otting G (2011) Suppression of isotope scrambling in cell-free protein synthesis by broadband inhibition of PLP enymes for selective 15N-labelling and production of perdeuterated proteins in H2O. J Biomol NMR. 50:35-42

32. Yokoyama J, Matsuda T, Koshiba S, Tochio N, Kigawa T (2011) A practical method for cell-free protein synthesis to avoid stable isotope scrambling and dilution. Anal Biochem 411:223-9. doi:10.1016/j.ab.2011.01.017

33. Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR Nature 445:618-622.

34. Macek P, Kerfah R, Boeri Erba E, Crublet E, Moriscot C, Schoehn G, Amero C, Boisbouvier J (2017) Unraveling Self-Assembly Pathways of the 468 kDa Proteolytic Machine TET2. Science Adv. doi: 10.1126/sciadv.1601601.

35. Mas G, Guan J-Y, Crublet E, Colas Debled E, Moriscot C, Gans P, Schoehn G, Macek P, Schanda P, Boisbouvier J (2018) Structural Investigation of a Chaperonin in Action Reveals How Nucleotide Binding Regulates the Functional Cycle. Science Adv. doi: 10.1126/sciadv.aau4196.

36. Franzetti B, Schoehn G, Hernandez JF, Jaquinod M, Ruigrok RW, Zaccai G (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea EMBO J. 21:2132-2138.

37. L. Borissenko, M. Groll, Crystal Structure of TET Protease Reveals Complementary Protein Degradation Pathways in Prokaryotes. *J. Mol. Biol.* **346**, 1207–1219 (2005).

38. Fraga H, Arnaud CA, Gauto DF, Audin MJC, Kurauskas V, Macek P, Krichel C, Guan JY, Boisbouvier J, Sprangers R, Breyton C, Schanda P (2017) Solid-state NMR H-N-(C)-H and H-N-C-C 3D/4D correlation experiments for resonance assignment of large proteins. Chem. Phys. Chem. 18:2697 –2703

39. Gauto D, Estrozi L, Schwieters C, Effantin G, Macek P, Sounier R, Sivertsen AC, Schmidt E, Kerfah R, Mas G, Colletier JP, Güntert P, Favier A, Schoehn G, Schanda P, Boisbouvier J (2019) Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. Nature Communication doi : 10.1038/s41467-019-10490-9.

40. Torizawa T, Shimizu M, Taoka M, Miyano H, Kainosho M. (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. J Biomol NMR 30:311-25
41. Apponyi MA, Ozawa K, Dixon NE, Otting G. (2008) Cell-free protein synthesis for analysis by NMR spectroscopy. Methods Mol Biol 426:257-68. doi: 10.1007/978-1-60327-058-8_16
42. Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A, Yokoyama S (2004) Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. J Struct Funct Genomics 5:63-68
43. Pedersen A, Hellberg K, Enberg J, Karlson G (2011) Rational improvement of cell-free protein synthesis. New Biotechnology 28:218-24. doi:10.1016/j.nbt.2010.06.015
44. Mayerhofer H, Sautron E, Rolland N, Catty P, Seigneurin-Berny D, Pebay-Peyroula E, Ravaud S (2016) Structural Insights into the Nucleotide-Binding Domains of the P1B-type ATPases HMA6 and HMA8 from Arabidopsis thaliana. PLoS One doi: 10.1371/journal.pone.0165666
45. Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. J Magn Reson 187:163-169
46. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 6:277-293.
47. Bertini I, Luchinat C, Parigi G, Ravera E, Reif B, Turano P (2011) Solid-state NMR of proteins sedimented by ultracentrifugation. Proc Natl Acad Sci U S A. 108:10396-10399. doi: 10.1073/pnas.1103854108.

Table 1. Preparation of reaction mix and feeding mix for CECF *in vitro* synthesis of protein

Amino acid mix	**Volume (µL)**
Water soluble AA: A, G, H, K, P, R, S, V, T 50 mM concentration for each AA	247.5
In 1 M HCl : C, D, E, L, M, N, Q, W, Y 50 mM concentration for each AA	247.5
In 1 M NaOH: I, F 50 mM concentration for each AA	247.5
RNAse free H_2O	82.5
TOTAL	**825.0**
10x reaction mix	**(µL)**
100 mM rCTP	96.0
100 mM rGTP	96.0
100 mM rUTP	96.0
2.0 M HEPES-KOH pH 7.5	330.0
100 mM ATP	144.0
10 mM Folinic Acid	81.6
100 mM cyclic AMP	76.8
1 M DTT	40.8
1M spermidine	24.0
9.2 M NH_4OAc	35.9
RNAse free H_2O	178.9

TOTAL	**1200**	
Reaction Mix (RM) and Feeding Mix (FM)	**RM (µl)**	**FM (µl)**
10X reaction mix	100.0	1000.0
1M Creatine Phosphate	80.0	800.0
Amino acid mix	66.7	666.7
4 M Potassium Glutamate	52.0	520.0
1.07 M $Mg(OAc)_2$	7.9	130.8
17.5 mg/ml MRE 600 tRNA	10.0	0
10 mg/ml Creatine Kinase	25.0	0
T7 RNA polymerase (1/100e)	10.0	0
Adjust pH of FM to 7.5 with KOH		
S30 extract	400.0	0
Target DNA (16 µg/mL)	16.0	0
RNAse free H_2O	232.5	6882.5
Volume (µL) of mix	**1000**	**10000**

Figure Legends :

Figure 1: Schematic illustration of strategies used to produce perdeuterated proteins *in vivo* or *in vitro*. (A) *In vivo* production of perdeuterated proteins in M9/$^{2}H_2O$ media. Purified proteins need to be destabilized in $^{1}H_2O$ to allow back-protonation of amide sites before refolding. (B) *In vivo* production of perdeuterated proteins in M9/$^{1}H_2O$ media supplemented with U-[^{2}H, ^{15}N] or U-[^{2}H, ^{13}C, ^{15}N]-amino acids to ensure complete protonation of amide sites. α−sites are partially protonated because of the residual transaminases activities. (C) Cell-free production of perdeuterated proteins in M9/$^{1}H_2O$ media supplemented with perdeuterated amino acids using $NaBH_4$ treated S30 cell extracts in the presence of transaminase inhibitors (DM, AOA) to suppress residual protonation on α−positions and ensure full protonation of amide sites.

Figure 2. Residual protonation levels (%) detected in Cα-position of proteins expressed *in vitro* in $^{1}H_2O$ using perdeuterated amino mix. For quantification purpose the Nucleotide-Binding Domain of the P1B-type ATPase HMA8 from *Arabidopsis thaliana* (16.7 kDa, 155 amino acids [44], was expressed as a model protein. Expression *in vitro* was undertaken in $^{1}H_2O$ following the protocols presented in section 3 to produce perdeuterated U-[^{2}H, ^{13}C, ^{15}N] protein using the batch mode (**A**) (*see* **Note 17**) or the CECF mode (**B**) without inhibitors (black) or with inhibitors (AOA, D-malate and $NaBH_4$ S30 extract treatment) (grey). In (A) hydrolyzed U-[^{2}H, ^{13}C, ^{15}N] algal extract (Celtone® *see* **Note 16**) was complemented with unlabeled tryptophan. In (B) cell-free mixture containing 20 U-[^{2}H, ^{13}C, ^{15}N] amino acids (Merck) was used. Quantification of residual protonation was obtained from 3D HNCA experiments [45] using the integration module of the NMRPipe software [46]. To distinguish both isotopomers, the deuterium was decoupled during ^{13}C edition, while the ^{1}H decoupling was omitted. Percentage of ^{1}H or ^{2}H isotopes on Cα positions was deduced from the volume ratio of the correlations corresponding to the ^{13}C-^{1}H and the ^{13}C-^{2}H pairs.

For each residue type, the results displayed are the percentage of residual protonation calculated from the mean of 2 to 7 amino acids (exceptions are for histidine in both production modes, phenylalanine and tryptophan in batch mode and proline in CECF mode, where only one amino acid was quantifiable). Quantification was not undertaken for methionine and aspartic acid because correlations were overlapping. Labeled tryptophan is absent in the sample produced in CECF mode (**B**) because it was added to the algal extract in an unlabeled form. The data for those amino acids, denoted by an asterix (*), were therefore taken from Otting and coll. [31]. There are no quantifications for cysteine as this protein does not have any in its sequence.

Figure 3: 2D-(^{15}N, ^{1}H) solid-state NMR spectra of protein TET2 acquired using a 600 MHz Bruker Avance III HD spectrometer equipped with a 1.3 mm triple-resonance MAS probe. U-[^{2}H,^{15}N] TET2 sample used to acquire the spectrum was produced *in vitro* (CECF mode), using $^{1}H_2O$ as a solvent during cell-free protein synthesis and amino acids mix from hydrolyzed U-[^{15}N,^{2}H] Celtone® complemented with unlabeled W and C amino acids. Sample was sedimented overnight at 65,000 *g* in a 1.3 mm rotor [47]. Heteronuclear transfers were performed using 750 ms cross polarization sequence (^{1}H: 15 kHz, ^{15}N: 40 kHz), the MAS frequency was set to 55 kHz and the effective temperature was 28°C [39].

Figure 4: 2D-(^{15}N,^{1}H) extracts from solution NMR spectra of the 468 kDa TET2 protein assembly using CRINEPT-HMQC-TROSY experiment [12] acquired on a 950 MHz Bruker Avance III HD spectrometer equipped with a 5 mm cryogenically cooled pulsed-field-gradient triple-resonance probe at a temperature of 50°C. U-[^{2}H,^{15}N] TET2 assemblies were produced *in vitro* (CECF mode), concentrated (200 µL at 10 µM) in 20 mM TRIS buffer (pH 7.4) containing 20 mM NaCl, and loaded in 4 mm shigemi tubes. In (A) the sample used to acquire spectra presented was produced using $^{1}H_2O$ as a solvent during Cell-free protein synthesis and amino mix from hydrolyzed U-

[^{15}N,^{2}H] Celtone® complemented with unlabeled W and C amino acids. Sample (B) was produced using the same amino acids source but $^{2}H_2O$ as a solvent during cell-free synthesis, and the purified protein was then dialyzed in $^{1}H_2O$ buffer and stored at room temperature for 10 days to allow ^{2}H->^{1}H substitution at exchangeable positions. Sample (C) was produced as described for sample (A), but the purified protein was dialyzed in $^{2}H_2O$ buffer and stored at room temperature for 10 days to allow ^{1}H->^{2}H substitution in order to detect signals of solvent protected amide positions only.

Figure 1

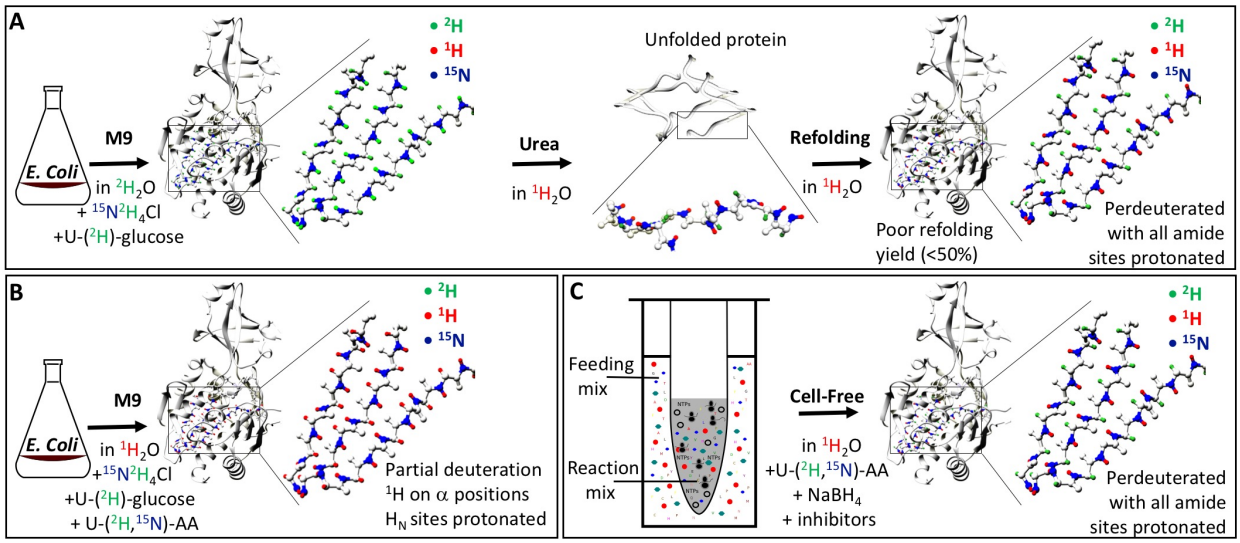
A
E. Coli
M9
in $^{2}H_2O$
+ $^{15}N^{2}H_4Cl$
+U-(^{2}H)-glucose
^{2}H
^{1}H
^{15}N
Urea
in $^{1}H_2O$
Unfolded protein
Refolding
in $^{1}H_2O$
Poor refolding
yield (<50%)
Perdeuterated
with all amide
sites protonated
B
E. Coli
M9
in $^{1}H_2O$
+$^{15}N^{2}H_4Cl$
+U-(^{2}H)-glucose
+ U-(^{2}H,^{15}N)-AA
Partial deuteration
^{1}H on α positions
H_N sites protonated
C
Feeding
mix
Reaction
mix
Cell-Free
in $^{1}H_2O$
+U-(^{2}H,^{15}N)-AA
+ $NaBH_4$
+ inhibitors
Perdeuterated
with all amide
sites protonated

Figure 2:

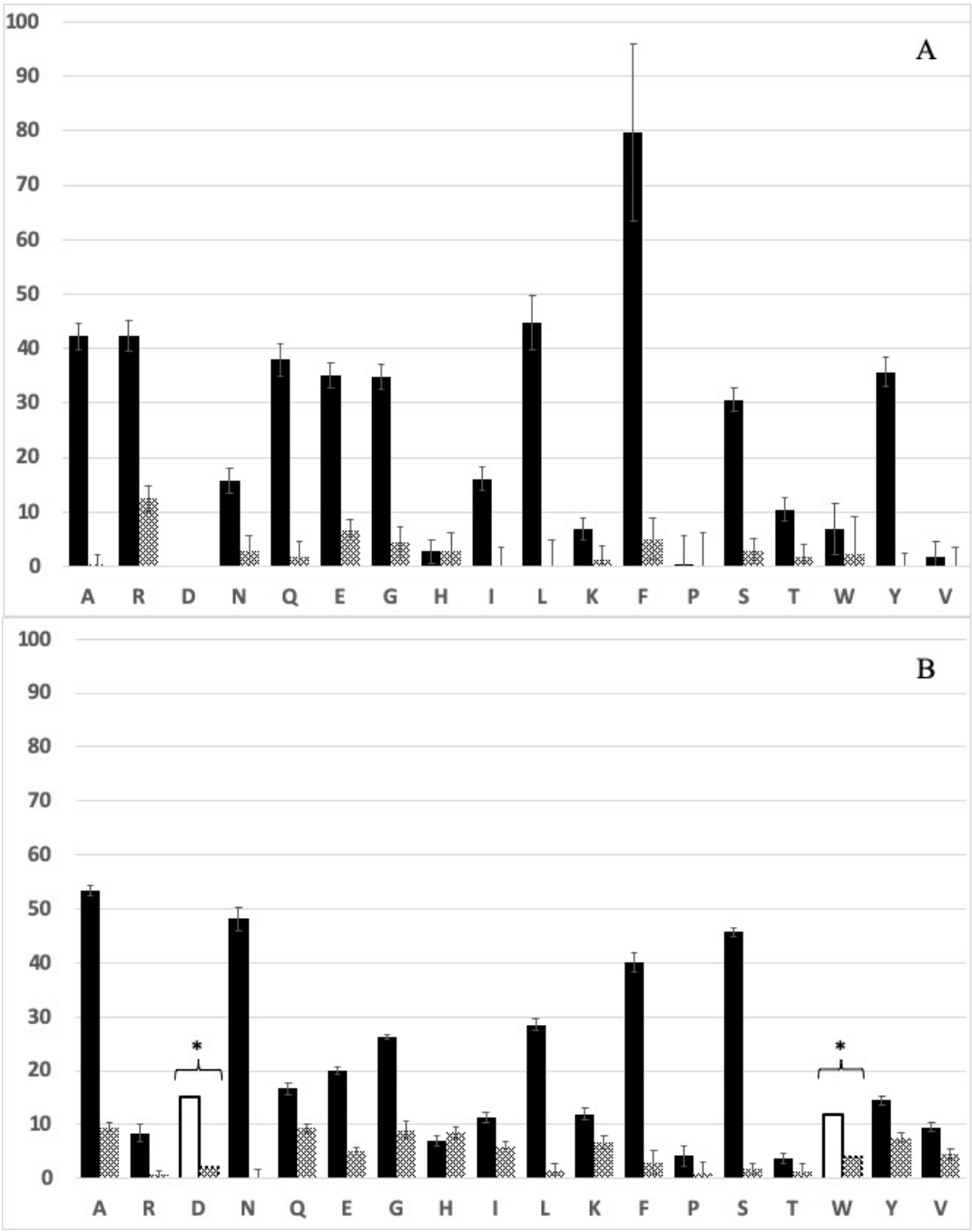

Figure 3:

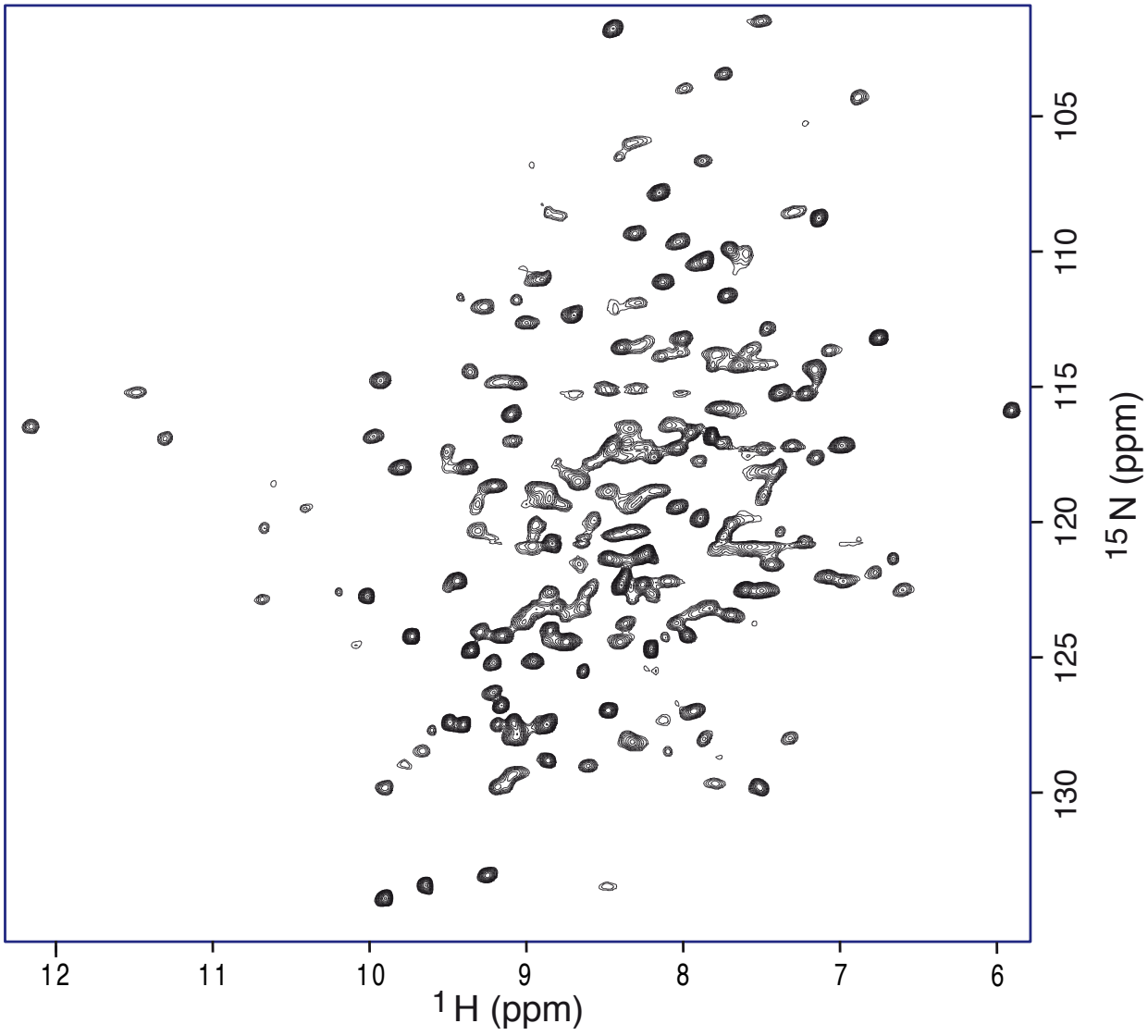

Figure 4: