
NMR Sample Preparation

3.1. INTRODUCTION

On many occasions, while carrying out interesting NMR projects, researchers have found that protein sample preparation is the bottleneck and most time-consuming stage of the planned studies. A similar situation occurs in crystallographic studies, however, a difference is that NMR samples usually require isotope labeling. Because the cost of ^{13}C , ^{15}N , or ^2H source compounds is significantly higher than natural abundant sources, the isotopic labeling of the proteins is usually done in minimal growth media using bacterial expression systems. This chapter will describe the common steps of protein sample preparation for NMR studies. Most of these steps are common for regular protein preparation by molecular biologists and biochemists. Some steps are especially NMR-oriented and will be emphasized in more detail. Some tips for sample preparation are also provided. It should be noted that there is often a great deal of flexibility in the application of protocols. Hence, it is often possible to alter or adapt a technique to specific needs.

Questions to be addressed in the present chapter include:

1. How do you choose bacterial expression systems for high expression of target proteins?
2. How are protein expression and solubility optimized?
3. What are minimal media?
4. How much ^{13}C and ^{15}N source compounds are needed to obtain sufficient isotope-labeled samples for heteronuclear NMR experiments?
5. What are the brief steps in protein purification for NMR samples?
6. What buffers, protein concentration, pH, and temperature are suitable for NMR samples?
7. What is a typical procedure for NMR sample preparation?
8. How are alignment media prepared for residual dipolar coupling measurement?
9. How is a protein–peptide complex sample prepared?
10. How is a protein–protein complex sample prepared?
11. Examples of NMR sample preparation (complete protocol).

3.2. EXPRESSION SYSTEMS

3.2.1. *Escherichia coli* Expression Systems

NMR samples normally require large quantities of isotope-labeled proteins or protein complexes at millimolar concentrations (can be less if a cryogenic probe or higher field spectrometers, such as 800 or 900 MHz spectrometers, are used; see Chapter 2 on Instrumentation). However, certain proteins or domains prepared using bacterial expression systems may have the following problems: (a) low expression level; (b) insolubility at high concentration; (c) low stability. These problems require large efforts of exploration and optimization of various conditions. The first step is typically to develop good expression systems using recombinant genetic techniques. The recombinant technology can yield substantially high levels of target proteins compared to the low level of proteins from the natural source. Moreover, recombinant expression systems can be manipulated to produce protein domains or to attach tags for easy purification. The former is particularly important for NMR studies since many NMR projects focus on the structures and dynamics of protein domains or domain–domain complexes. The common steps for protein subcloning can be found in many molecular biology books (such as *Short Protocols in Molecular Biology* by Ausubel *et al.*, 1997) and will not be detailed here. What we will focus on here are the steps in choosing the expression vectors for subcloning.

Compared to insect or mammalian cell expression systems, the prokaryotic *E. coli* expression systems are most commonly used to produce isotope-labeled proteins for NMR studies. One can start with many commercial *E. coli* expression vectors available, such as the pET system from Novagen, Inc., the Impact system from New England Biolabs, etc. (Table 3.1), which are routinely used in most NMR laboratories. Because individual proteins may behave differently in different expression systems, it is frequently difficult to predict the best expression vector for a particular protein. For initial investigation, one usually starts with vectors encoding protease-cleavable N-terminal or C-terminal fusions such as His-tag, GST, etc., which allows easy purification using an affinity column (see section 3.4). A His-tag is a flexible linker containing only 6–20 residues. Hence, it is usually uncleaved after affinity purification since its short size imposes minimal complication of the NMR spectra of fogged proteins. For example, Figure 3.1 (see color insert after Chapter 8) shows the two-dimensional ^1H - ^{15}N HSQC of a fragment from the cytoskeletal protein talin with and without a His-tag. One can

TABLE 3.1
Common Expression Vectors

Vector name	Fusion partner	Source	Features
pET vectors	His-tag or none	Novagen, Inc.	Short His-tag allows quick NMR screening of uncleaved protein
pGEX vectors	GST	Novagen, Inc.	Easy for purification but GST is a dimer that may interfere with target protein folding
pMal vectors	MBP	New England Biolabs, Inc.	Highly soluble MBP promotes expression and solubility of the target proteins but MBP has a low affinity for the resin, which reduces the yield
pTYB11 vectors	Intein	New England Biolabs, Inc.	Protein purified without fusion attached

immediately conclude that the fragment is folded and the signals are well dispersed, which means that the protein is suitable for further investigation. It is clear that although the spectrum has an improved resolution after cleavage, the signals of those unstructured His-tag residues do not detract from the feasibility of the project.

3.2.2. Fusion Proteins in the Expression Vectors

In many cases, different fusion proteins lead to substantially different expression levels and solubility of the target proteins. Thioredoxin, protein B G1 domain, and maltose-binding protein (MBP) are known to promote high expression as well as high solubility of the target proteins. Protein B G1 domain is small (~56 residues) and does not cause severe complication of spectral analysis and, hence, can be used as a sample solubility enhancer of the target proteins or protein complexes. However, caution must be paid since fusion proteins may prevent folding or induce misfolding of the target proteins, particularly when the N-terminal regions of the proteins are important interior components of protein structures. In these cases, proteins are usually found in insoluble inclusion bodies or look unfolded as judged from the NMR spectra. To avoid this problem, vectors without fusions such as pET3a (Table 3.1) can be exploited.

3.2.3. Optimization of Protein Expression

When proteins or domains have low expression level and low solubility, a number of expression vectors need to be explored for systematic and sometimes time-consuming optimization. The first thing to do to increase protein expression is to choose appropriate cell lines. The most common cell lines (strains) for bacterial protein expression are BL21(DE3), BL21(DE3) pLYS, HMS, etc., which are commercially available. The cDNA containing the target protein is usually transformed into the above strains on Day 1. The next day, one colony is picked from each transformation and grown in a 5 ml culture for each strain. Cell density (OD_{600nm}) is checked after a few hours and 1 mM IPTG (isopropyl-1-thio- β -D-galactoside) is added (typical for the first time but can be varied, see below) when the OD is approximately 0.6 to induce protein expression (1 ml of culture is sampled before addition of IPTG). The cells are harvested after 3–4 hr and spun down at 10,000 g. The pellets are lysed by sonication using a standard PBS (Phosphate-Buffered Saline) buffer (100 μ l). To check the protein expression, 2 μ l of the lysate from each pellet (before and after IPTG) are run on an SDS PAGE gel. This will give some idea of which strain gives the highest expression. The next step is to check protein solubility by spinning down the lysate containing IPTG and taking the supernatant to check if there is soluble protein. Hence, contained in the SDS PAGE will normally be three lanes for each strain: Lane 1, uninduced lysate; Lane 2, lysate induced with IPTG; Lane 3, supernatant for the induced lysate. If proteins are expressed in inclusion bodies, refolding protocols may be used (see section 3.4) but many proteins cannot undergo the reversible unfolding/refolding process. To maximize protein expression and to increase the soluble fraction in bacterial lysates, one usually needs to vary a series of expression conditions:

1. Prepare a cell culture growth curve to decide the best induction point. Each cell line may have a different growth curve. The growth curve can be made by measuring cell density at 600 nm as a function of time in hours.

2. Induce protein expression at the middle of the log-phase derived from step 1 with different IPTG concentrations typically from 0.1 mM to 2 mM.
3. Vary the induction time between 1 hr and overnight at different temperatures. Lower temperatures such as 16°C are often useful to produce more soluble proteins than higher temperatures such as 37°C.

Sometimes, protein solubility is still low after the above procedures due to inappropriate conditions for protein folding in bacteria. In such a case, insect and mammalian cell expression systems may be used to assist the native folding of the proteins, although the cost of isotope labeling will be extremely expensive. These methods are not detailed here.

3.3. OVEREXPRESSION OF ISOTOPE-LABELED PROTEINS

After optimization of protein expression, the next step is to overexpress in large scale isotope-labeled proteins for NMR studies. Heteronuclear multidimensional NMR experiments for structure determination of medium-sized protein requires that the target proteins or domains be uniformly ^{15}N and/or ^{13}C -labeled. This is done by growing cell cultures in minimal media in which $^{15}\text{NH}_4\text{Cl}$ and/or ^{13}C glucose are the only sources for nitrogen and carbon atoms. The standard recipe for the minimal media is shown in Table 3.2 and the recipe for Trace Element Solution in Table 3.3. The most common cell lines for protein expression in minimal media are BL21(DE3), BL21(DE3) pLYS, HMS, etc., which are commercially available. Note that protein expression levels are usually lower in minimal media as compared to rich LB media. Hence, it is recommended that different cell lines be used to explore optimum expression. Due to the expensive cost of ^{13}C glucose, it is recommended to optimize glucose usage at small scale by measuring the growth curve as a function of different amounts of unlabeled glucose. Protein expression levels should be checked and compared on SDS PAGE by taking 1 ml of culture before and after the IPTG induction. Although a typical induction time for protein

TABLE 3.2
Recipe for Minimal Media^a

Compound	Amount	Comments
K_2HPO_4	10.0 g/L	
KH_2PO_4	13.0 g/L	
Na_2HPO_4	9.0 g/L	
K_2SO_4	2.4 g/L	
$^{15}\text{NH}_4\text{Cl}$	1–2 g/L	~\$40/g
^{13}C Glucose	2–5 g/L	Amount variable and needs to be optimized to reduce the cost (~\$130/g)
Trace element solution ^b	10 ml/L	See Table 3.3
1 M $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$	10 ml/L	
Thiamine (Vitamin B ₁)	5 mg/ml	
Antibiotics	~0.1 mg/L	

^a The media have to be sterilized.

^b The trace element solution is a combination of trace elements shown in Table 3.3.

TABLE 3.3
Recipe for Trace Element Solution^a

Compound	Grams/100 ml ^b
CaCl ₂ ·2H ₂ O	0.600
FeSO ₄ ·7H ₂ O	0.600
MnCl ₂ ·4H ₂ O	0.115
CoCl ₂ ·6H ₂ O	0.080
ZnSO ₄ ·7H ₂ O	0.070
CuCl ₂ ·2H ₂ O	0.030
H ₃ BO ₃	0.002
(NH ₄) ₆ Mo ₇ O ₂₄ ·4H ₂ O	0.025
EDTA	0.500

^a Add ingredients one at a time, waiting 5–10 min before they fully dissolve. After adding EDTA and stirring for a few hours, the color of the solution should be golden brown (if it is greenish, then leave stirring overnight). Sterilize afterwards by filtering through a 0.2 μ M filter.

^b A fresh stock of 100–200 ml is usually made.

expression is 2–4 h, induction time is 2–3 times longer when the proteins are expressed in partial/full ²H₂O due to the slow growth rate. Growing cultures in ²H₂O is required to prepare deuterium labeled proteins for NMR studies when the proteins are relatively large (>20 kDa).

3.4. PURIFICATION OF ISOTOPE-LABELED PROTEINS

Purification of isotope-labeled proteins is a key step and probably the most time-consuming step for NMR sample preparation. The procedures and tips for purifying the isotope-labeled proteins are the same as for non-labeled proteins described in many textbooks and the literature. If the labeled proteins contain fusions, fusion-targeted affinity columns will be the first step for purification followed by protease cleavage and gel filtration. This procedure typically works if the protein behaves during the process; however, protease cleavage sometimes can be a tricky process. Excess amounts of protease or over-digestion by protease can lead to non-specific cleavage and hence optimization is usually required. When proteins are not fused, the chemical structure and physical properties of the proteins are the two key parameters used to develop the most efficient purification protocols. Isoelectric point (pI), pH stability, and charge density are important properties to be exploited during purification. Several steps of different ion-exchange and hydrophobic chromatography are often used for large scale purification of non-fused proteins followed by a final step of gel-filtration. Note that 90% pure proteins are usually sufficient for heteronuclear NMR studies if the proteins are stable in the presence of the impurities. Some important tips for purification are summarized below:

1. A French press is often better than sonication in lysing the cells by producing more soluble fractions of proteins.
2. A cocktail of protease inhibitors (Table 3.4) is recommended in the cell lysis buffer. A typical lysis buffer for soluble proteins consists of 20–50 mM phosphate, pH 7.4, or 0.1 M Tris-HCl, pH 7.4, 0.1 M NaCl, 1–5 mM EDTA (assuming the protein of

TABLE 3.4
Cocktail of Protease Inhibitors

Inhibitor	Target protease	Final concentration ($\mu\text{g/ml}$)
Leupeptin	Broad spectrum	0.5–10
EDTA-Na2	Metalloproteases	5–10
Pepstatin A	Acidic proteases	0.7–10
Aprotinin	Serine proteases	50
PMSF	Serine proteases	0.2–2
Benzamidine HCl	Serine proteases	100
Soybean	Trypsin-like trypsin inhibitor	100

interest has no metal), 5–20 mM β -mercaptoethanol (assuming the protein of interest has no disulfide bonds), sucrose and the cocktail of protease inhibitors. Note that the conditions of lysis buffer vary significantly depending on the properties of the proteins. For example, low salt or no salt is used for Tris buffer if the protein or domain of interest behaves poorly in the presence of salt. The yield can differ by 2–10-fold between two different buffer conditions.

3. Try different purification protocols to optimize the yield. To develop the best purification protocol, one should always try different protocols for initial screening. For example, if most of the proteins go into inclusion bodies, it may be beneficial to try a refolding protocol. As mentioned above, some proteins are not reversible in folding/unfolding, but some proteins may be refoldable.

3.5. NMR SAMPLE PREPARATION

3.5.1. General Considerations

The last step for NMR sample purification is to choose a good buffer in which the protein is concentrated to approximately 1 mM. Phosphate buffer at pH 5–7 (20–50 mM) with or without salt (e.g., KCl, NaCl) is often used for many NMR samples. It is recommended to try a series of conditions in small scales and then decide which condition is the best. Quite frequently, a protein/domain itself is not very stable in the buffer but becomes very stable after mixing with the target protein/peptide. High quality NMR tubes with appreciate specifications should be used for protein samples, which are usually tubes 5 mm in diameter containing 0.5 ml 95% H_2O /5% $^2\text{H}_2\text{O}$ for aqueous samples. If the volume of the sample is limited, micro tubes for which the susceptibility matches that of $^2\text{H}_2\text{O}$ are chosen for a total sample volume of approximately 200 μl , such as Shigemi micro tubes (Shigemi Inc., Allison Park, PA). Because of the small sample volume, the buffer contains 7% $^2\text{H}_2\text{O}$ used for ^2H lock. In addition, the samples are usually required to be degassed by blowing high purity argon or nitrogen gas into them to remove oxygen—the paramagnetic property of which will broaden the line shapes of protein resonances.

3.5.2. Preparation of Protein–Peptide Complexes

The contact surface contributing to the interactions of high affinity and specificity often involves 30 or less amino acid residues from each protein of the complex (de Vos *et al.*, 1992;

Song and Ni, 1998). Often this contact surface is located in a single continuous fragment of one of the proteins, which can be identified by mutation and deletion experiments. Therefore, fragments can be chemically synthesized in large amounts and studied by two-dimensional ^1H NMR experiments due to their small molecular size (Wüthrich, 1986). Samples for protein–peptide complexes are commonly prepared from isotopically labeled protein and unlabeled peptide according to the following procedure, since the availability of labeled peptide is often prohibited by the expense of chemical synthesis from labeled amino acids and the difficulty of biosynthesis due to peptide instability during its expression and purification (Huth *et al.*, 1997; Newlon *et al.*, 1997).

Preparation of the complexes is done by titrating synthetic peptide (unlabeled) into the isotope-labeled target protein (Breeze, 2000; Qin *et al.*, 2001). The stoichiometry of association can be best determined by monitoring a ^1H - ^{15}N HSQC spectrum of the target protein in different protein/peptide ratios. Because of the high sensitivity of the ^1H - ^{15}N HSQC experiment, only a low concentration of ^{15}N labeled protein (0.1–0.2 mM) is needed for the titration experiments. Once the stoichiometry is determined, the labeled protein is preferably mixed with unlabeled peptide at dilute concentration and then concentrated to 0.5–1 mM, required for most NMR experiments. A higher field magnet or high sensitivity probe such as a cryogenic probe makes it possible to study dilute samples if aggregation occurs at high concentrations. The pH of the samples is preferably kept below 7.0 to reduce the amide exchange rate. Further purification of the tightly associated complex by gel-filtration may be necessary to improve the sample quality.

3.5.3. Preparation of Protein–Protein Complexes

Protein–protein interactions play an essential role at various levels in information flow associated with various biological processes, such as gene transcription and translation, cell growth and differentiation, neurotransmission, and immune response. The interactions frequently lead to changes in the shape or dynamics as well as the chemical or physical properties of proteins involved. Solution NMR spectroscopy provides a powerful tool to characterize these interactions at the atomic level and at near physiological conditions. With the use of isotopic labeling, the structures of many protein complexes in the 40 kDa total molecular mass regime have to be determined (Clore and Gronenborn, 1998). The development of novel NMR techniques and sample preparation has been further increasing the mass size available for the structural determination of protein complexes. Furthermore, NMR has been utilized to quickly identify the binding sites of the complexes based on the results of chemical shift mapping or hydrogen-bonding experiments. Because it is particularly difficult and sometimes impossible to crystallize weakly bound protein complexes ($K_d > 10^{-6}$), the chemical mapping method is uniquely suitable to characterize such complexes. The binding surfaces of small to medium sized isotopically labeled proteins with molecular mass less than 30 kDa to large target proteins (unlabeled, up to 100 kDa) can be identified by solution NMR (Mastaglio *et al.*, 1999; Takahashi *et al.*, 2000). As discussed in Chapter 6, the structure of small ligands weakly bound to the proteins can be determined by transferred NOE experiments (Clore and Gronenborn, 1982, 1983). The structures of the peptides or small protein domains of weakly bound protein complexes can also be characterized by the NMR technique, which may be beneficial to the discovery and design of new drugs with high affinity. In addition to the structural investigation of protein complexes, NMR is a unique and powerful technique to study the molecular dynamics involved in protein–protein reorganization (Kay *et al.*, 1998; Feher and Cavanagh, 1999). Furthermore, protein binding sites often contain residues from different parts of the protein

or domain. Structure determination of protein–protein (domain) complexes is necessary for understanding specificity.

Preparation of protein–protein complexes (Breeze, 2000; Qin *et al.*, 2001): for a complex A–B, the A and B components are separately expressed and purified. The isotope-labeled A is mixed with unlabeled B or vice versa to simplify the NMR spectra. Because of the large sizes of complexes, partial or full deuteration may be necessary to reduce the line widths of signals. Isolated domains are sometimes unfolded or partially folded and may also undergo aggregation, which makes purification difficult. Purification can be performed in the denatured condition and the unfolded domains can be refolded in the presence of target proteins. Fusion proteins are often needed to help for solubilization of domains for purification. Fusion proteins are cleaved after the domain is mixed with the target protein for stabilization. The complex is further purified by gel-filtration that removes impurities including the fusion protein and protease.

3.5.4. Preparation of Alignment Media for Residual Dipolar Coupling Measurement

Various media are available for moderately aligning macromolecules in solution in the magnetic field, of which two liquid crystalline media are most commonly used at the present time: DMPC/DHPC bicelles and filamentous phage. It is the interaction of the magnetic field with the anisotropic susceptibility of liquid crystalline media that aligns these particles in the magnetic field. When particles with a nonspherical shape, such as discs or rods dissolved in solution, are placed in the magnetic field, the anisotropic distribution of the electron density leads to an orientational dependence of this interaction. If the anisotropic interaction is large enough to overcome the thermal energy of the particles, the degree of orientation order of the media in the magnetic field becomes significant enough to be measurable, which is usually in the range of 0.5–0.85. As discussed previously (section 1.8.4 on Residual Dipolar Coupling), the order for macromolecules described by the magnitude of the alignment tensor is in the range of 10^{-3} , meaning that the interaction between aligned particles and macromolecules must be very weak. The weakness of the interaction is necessary so that it does not perturb the native structure of macromolecules under study or broaden the resonance line shape due to a high degree of order or a change of relaxation properties.

Disc-shaped bicelles (bilayered micelles) were the first liquid crystalline medium used to achieve weak alignment of macromolecules (Tjandra and Bax, 1997), and were originally developed by Prestegard, Sanders, and coworkers (Sanders and Prestegard, 1990, 1991; Sanders *et al.*, 1994). The medium contains a mixture of the saturated lipids dimyristoylphosphatidylcholine (DMPC) and dihexanoylphosphatidylcholine (DHPC) in low concentration to form planar bicelles in which DMPC constitutes the planar bilayer region and DHPC stabilizes the rim of the bicelles. The bicelles of the medium align in the presence of a magnetic field at 35°C whereas they are isotropic at 25°C.

The order of alignment media is transferred to macromolecules by rapid random collisions and electrostatic interaction between medium particles and the molecules dissolved in the medium. The formation of liquid crystals by the DMPC/DHPC mixture depends on a number of factors including the temperature, concentration, and molar ratio of the mixture, and ionic strength. The presence of other charged amphiphilic compounds can also influence the phase transition from isotropic to liquid crystalline phase. To maintain the weak interaction, the liquid crystalline medium is limited to a very dilute concentration, typically about 10% w/v. In such

dilute lipid concentration, the molar ratio of DMPC/DHPC plays an important role in the formation of bicelles. If the ratio is too low, the lipids form small size discs that are too small to generate measurable alignment. On the other hand, a high ratio causes the oversized bicelles to collapse to form spherical micelles by DMPC. The upper limit of the ratio at which stable planar bicelles can be formed is 5. Usually, the ratio is maintained in the range of 3.0–3.5, which corresponds to bicelle diameters of 200–250 Å with a thickness of approximately 40 Å.

At the low concentration, the DMPC/DHPC bicelles are unstable in the lower range (25–30°C) of liquid crystalline phase temperature (25–45°C). The stability of the bicelles can be improved by adding a small amount of charged amphiphile (Losonczi and Prestegard, 1998), such as cetyl (hexadecyl) trimethyl ammonium bromide (CTAB, positively charged) or sodium dodecyl sulfate (SDS, negatively charged). Best results are obtained with a molar ratio of 0.01–0.003 relative to DMPC (e.g., DMPC : DHPC : CTAB = 3.5 : 1.0 : 0.005). The addition of these detergents widens the temperature range of liquid crystalline media, resulting in stabilizing the bicelles in the above original temperature range. Charging the bicelles with the detergents also produces an electrostatic potential that attracts and repels groups with opposite and like charges, respectively, resulting in a change in orientation and magnitude of the alignment tensor. This change is often enough to yield an independent alignment tensor and reduce the ambiguity of orientation mentioned early.

The procedure for the preparation of bicelle samples includes a number of straightforward steps (Ramirez *et al.*, 2000). It starts with weighing the appropriate amount of dry powder DMPC and DHPC according to the desired molar ratio, typically 3.5 : 1. The weighing should be done in a dry box to obtain an accurate amount of the material due to the hygroscopic property of DHPC. The buffer is prepared with the required amount of salt and sodium azide at a concentration greater than or equal to 1 mM as an antibacterial agent, which is then added into DHPC. Dry powder DMPC is then added into the sample. In order to completely dissolve DMPC, cooling and heating cycles are repeated until the solution becomes clear by freezing the solution in liquid nitrogen and thawing it while vortexing at 35°C. An alternative method is to leave the mixture solution at 20°C for a few days to let it turn into a clear solution. It should be noted that a protein sample cannot be recovered from the liquid crystalline medium once it is dissolved in it. The alignment media may be frozen or lyophilized (Bax *et al.*, 2001).

Filamentous phage is also utilized to achieve weak alignment of macromolecules in the magnetic field (Clore *et al.*, 1998; Hansen *et al.*, 1998a,b) and is commercially available (Asla Labs). Because of their rod shape (1–2 µm long, 6.5 nm diameter), filamentous phage with a certain concentration in solution (as low as a few mg/ml) can form liquid crystals. When the sample is placed in the magnetic field, the particles are aligned with their long axis parallel to the field direction due to their very anisotropic nature. In order to obtain the liquid crystalline phase, the solution should be prepared at a pH higher than neutral to maintain a negatively charged environment so as to prevent glutamate and aspartyl side chains of the phage-coated protein from protonating. In addition, a high salt concentration should be avoided at low phage concentration, because it can prevent alignment. Because of the relatively strong electrostatic interactions between phage and macromolecules relative to those between micelles and macromolecules, phage liquid crystallines generate different alignment tensors than in micelle media, resulting in additional alignment tensors for application to the reduction of orientational degeneracy. Phage samples are easily prepared by dissolving 3–10 mg/ml *pfl* (or *pd*) phage in the buffer solution. The liquid crystalline phase is formed over a wide range of temperatures (5–50°C). However, the phage medium cannot be frozen or lyophilized. Protein in the sample is usually recovered by precipitating the phage through centrifugation.

3.6. EXAMPLES OF PROTOCOLS FOR PREPARING $^{15}\text{N}/^{13}\text{C}$ LABELED PROTEINS

3.6.1. Example 1: Sample Preparation of an LIM Domain Using Protease Cleavage

3.6.1.1. Background

PINCH LIM1 domain is a double zinc finger involved in cell adhesion. The protein was subcloned into several expression vectors including pET3a, pGEX-4T, and pMAL-C2X. Only pMAL-C2X gives good expression and, hence, was used for sample preparation.

3.6.1.2. Protein Expression

Expression plasmid pMAL-C2X, encoding an MBP fused to the N-terminus of residues 1–70 of human PINCH protein via a Factor X_a (FX_a) cleavable linker, was used for preparation of the NMR sample. Residues 1–70 contain the entire LIM1 domain. Due to cloning artifacts, the C-terminus of LIM1 had three additional residues (WIL), whereas the N-terminus contained four (ISEF). BL21 (DE3) cells harboring plasmid were grown in LB medium or in M9 minimal medium (Table 3.2) in the presence of 100 $\mu\text{g}/\text{ml}$ ampicillin. For isotope labeling, M9 contained 1.1 g/L $^{15}\text{N}\text{-NH}_4\text{Cl}$ and unlabeled or 3 g/L ^{13}C -labeled glucose. Three liters of culture were induced at OD_{600nm} approximately equal to 0.5 for 4 hr at 37°C with 1 mM IPTG.

3.6.1.3. Protein Purification and Sample Preparation

Cells were lysed with a French press and the cleared lysates were fractionated on a DEAE-sepharose column (50 mM Tris-HCl, pH = 8.0, gradient of NaCl 0.0–0.8 M). MBP-LIM1-containing fractions were concentrated, the buffer was exchanged into that optimal for cleavage (50 mM Tris-HCl, 100 mM NaCl, 3.5 mM CaCl₂, pH = 8.0) and subjected to FX_a treatment. Cleaved LIM1 was further purified on a Superdex 75 gel-filtration column. Fractions containing LIM1 were pooled and concentrated to approximately 0.5 mM with a buffer at pH = 7.5 containing 50 mM Na₂HPO₄, 100 mM NaCl, 0.5 mM β -mercaptoethanol.

3.6.2. Example 2: Sample Preparation Using a Denaturation–Renaturation Method

3.6.2.1. Background

The double-stranded RNA (dsRNA) activated protein kinase (PKR) contains a dsRNA binding domain (dsRBD), which was subcloned into pET15b vector with a thrombin-cleavable His-tag linker. However, the protein is expressed in inclusion bodies and, hence, the refolding method was used for protein purification.

3.6.2.2. Protein Expression

Four liters of *E. coli* BL21(DE3) pLYS cells transformed with pET15b encoding the dsRBD of human PKR were grown at 37°C in minimal media containing 0.4% glucose/0.1% $^{15}\text{NH}_4\text{Cl}$ or 0.4% [$^{13}\text{C}_6$] glucose/0.1% $^{15}\text{NH}_4\text{Cl}$ in order to obtain ^{15}N - and $^{15}\text{N}/^{13}\text{C}$ -labeled proteins, respectively. Cells were grown in log phase to OD_{600nm} = 0.6–0.8 in the presence of 50 $\mu\text{g}/\text{ml}$ carbenicillin, and protein expression was induced for 4 hr with 1 mM IPTG. The cells were harvested by centrifugation for 20 min at 6000 $\times g$ at 4°C, and the pellets were drained and stored at –80°C.

3.6.2.3. Protein Purification

The pellets were resuspended in ice-cold lysis buffer (6 M guanidine HCl and 50 mM Tris-HCl, pH 8.0), sonicated 4×30 s at full power, and centrifuged at $20,000 \times g$ for 20 min at 4°C. The supernatant was passed over an Ni^{2+} -agarose metal affinity column, and the histidine-tagged protein was eluted according to the manufacturer's instructions. The denatured dsRBD was further purified by gel filtration chromatography on a Superdex-75 column and refolded by dialysis against 50 mM phosphate buffer at 4°C, pH 6.5, and 1 mM DTT. The refolded dsRBD binds to dsRNA with the same affinity as wild type PKR and was concentrated by a Centrplus-10 ultrafiltration device (Amicon). The dsRBD is mostly monomeric (>90%) as judged by gel-filtration column, and its purity and concentration were determined by SDS-PAGE with Coomassie staining and UV spectrometry. The final ^{15}N -labeled or $^{15}\text{N}/^{13}\text{C}$ -labeled protein was purged with argon and adjusted to 7% $^2\text{H}_2\text{O}$, and transferred to 250 μl microcell NMR tubes (Shigemi Inc., Allison Park, PA) for NMR experiments.

QUESTIONS

- 3.1. Why are proteins needed in isotopic labeling for structure and dynamics studies?
- 3.2. Why are minimum media used to isotope label the proteins and what are the differences between minimum media and rich media?
- 3.3. What is His-tag and can it affect the quality of the NMR spectrum?
- 3.4. When is a micro NMR tube needed for NMR sample preparation?
- 3.5. What are the temperature ranges of the alignment media prepared using bicelles and filamentous phage?
- 3.6. For what purpose is a charged amphiphile such as CTAB added to the bicelles?

REFERENCES

- Ausubel, F., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith *et al.* (eds), *Short Protocols in Molecular Biology*, 3rd edn, Wiley & Sons, New York (1997).
- Bax, A., G. Kontaxis, and N. Tjandra, *Meth. Enzymol.* **339**, 127 (2001).
- Breeze, A., *Prog. Nucl. Magn. Reson. Spectrosc.* **36**, 323 (2000).
- Clore, G.M. and A.M. Gronenborn, *J. Magn. Reson.* **48**, 402 (1982).
- Clore, G.M. and A.M. Gronenborn, *J. Magn. Reson.* **53**, 423 (1983).
- Clore, G.M. and A.M. Gronenborn, *Trends Biotech.* **16**, 22 (1998).
- Clore, G.M., M.R. Starich, and A.M. Gronenborn, *J. Am. Chem. Soc.* **120**, 10571 (1998).
- de Vos, A.M., M. Ultsch, and A.A. Kossiakoff, *Science* **255**, 306 (1992).
- Fehér, V.A. and J. Cavanagh, *Nature* **400**, 289 (1999).
- Hansen, M.R., L. Mueller, and A. Paridi, *Nat. Struct. Biol.* **5**, 1065 (1998b).
- Hansen, M.R., M. Rance, and A. Paridi, *J. Am. Chem. Soc.* **120**, 11210 (1998a).
- Huth, J.R., C.A. Bewley, B.M. Jackson, A.G. Hinnebusch, G.M. Clore, and A.M. Gronenborn, *Prot. Sci.* **6**, 2359 (1997).

- Kay, L.E., D.R. Muhandiram, G. Wolf, S.E. Shoelson, and J.D. Forman-Kay, *Nat. Struct. Biol.* **5**, 15 (1998).
- Losonczi, J.A. and J.H. Prestegard, *J. Biomol. NMR* **12**, 447 (1998).
- Matsuo, H., K.J. Walters, K. Teruya, T. Tanaka, G.T. Gassner, S.J. Lippard *et al.*, *J. Am. Chem. Soc.* **121**, 9903 (1999).
- Newlon, M.G., M. Roy, Z.E. Hausken, J.D. Scott, and P.A. Jennings, *J. Biol. Chem.* **272**, 23637 (1997).
- Qin, J., O. Vinogradova, and A.M. Gronenborn, *Meth. Enzymol.* **339**, 377 (2001).
- Ramirez, B.E., O.N. Voloshin, R.D. Camerini-Otero and, A. Bax, *Prot. Sci.* **9**, 2161 (2000).
- Sanders, C.R. and J.H. Prestegard, *Biophys. J.* **58**, 447 (1990).
- Sanders, C.R. and J.H. Prestegard, *J. Am. Chem. Soc.* **113**, 1987 (1991).
- Sanders, C.R., B.J. Hare, K.P. Howard, and J.H. Prestegard, *Prog. Nucl. Magn. Reson. Spectrosc.* **26**, 421 (1994).
- Song, J. and F. Ni, *Biochem. Cell Biol.* **76**, 177 (1998).
- Takahashi, H., T. Nahanishi, K. Kami, Y. Arata, and I. Shimada, *Nat. Struct. Biol.* **7**, 220 (2000).
- Tjandra, T. and A. Bax, *Science* **278**, 1111 (1997).
- Wüthrich, K., *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York (1986).