Journal of Biomolecular NMR (2005) 31: 11–19

Quantitation of protein expression in a cell-free system: Efficient detection of yields and ^{19}F NMR to identify folded protein

Muniasamy Neerathilingam[a], Lesley H. Greene[b], Simon A. Colebrooke[a], Iain D. Campbell[a] & David Staunton[a,*]

[a]*Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.;* [b]*Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford OX1 3TA, U.K.*

Received 14 July 2004; Accepted 8 October 2004

Key words: cell-free protein expression, filter assay, 6-fluorotryptophan, ^{19}F NMR, high-throughput screening, retinol binding protein

Abstract

We have developed an efficient and novel filter assay method, involving radioactive labelling and imaging, to quantify the expression of soluble proteins from a cell-free translation system. Here this method is combined with the conformational sensitivity of ^{19}F NMR to monitor the folded state of the expressed protein. This report describes the optimisation of 6-fluorotryptophan incorporation in a His-tagged human serum retinol-binding protein (RBP), a disulphide bonded β-barrel protein. Appropriate reagent concentrations for producing fluorine labelled RBP in a cell-free translation system are described. It is shown that ^{19}F NMR is a suitable method for monitoring the production of correctly folded protein from a high-throughput expression system.

Introduction

The requirement for correctly folded, stable proteins with good solution behaviour is often the limiting factor in studies of protein structure and function (Staunton et al., 2003). Structural genomic projects have addressed this by introducing high-throughput (HTP) screening, where cDNA libraries are expressed and suitable samples for further study are identified, normally by expression levels and solubility. However, there is a relatively low success rate in going from construct to well-formed crystal or NMR sample. This is so even when the targets have been selected for stability (e.g., thermophiles) and potentially problematic targets have been excluded (e.g., integral membrane proteins) (Christendat et al., 2000). An obvious solution would be to optimise the conditions for a specific protein rather than relying on a generic set of conditions and to have a simple but reliable test for protein folding.

Cell-free translation is ideally suited for HTP protein expression due to the open nature of the system and it is being successfully used in structural genomics projects (Yokoyama, 2003). The main attraction of cell-free expression for high-throughput screening is that constructs and conditions can be explored in batch-mode reactions of less than 100 μl, allowing a 96-well format to be used. In this format the yields of both soluble and insoluble protein are followed by radioactive, immunological or fluorescent labels until optimized (Busso et al., 2003). The successful batch reactions can then be scaled up to the millilitre level in a semi-continuous batch reaction in

*To whom correspondence should be addressed: E-mail: staunton@bioch.ox.ac.uk

which the expression system is supplied with reagents from a reservoir through a dialysis membrane (Kigawa and Yokoyama, 1991). This extends the life of the reaction to more than 12 h and the protein yield to the milligrams required for sample preparation (Kigawa et al., 1999, 2001). Another advantage is that the amino acids are chosen by the researcher and therefore labelling can be specific to particular amino acids (Kigawa et al., 1995, 2001). Artificially charged suppressor tRNAs can also be used to label specific positions in the protein sequence (Yabuki et al., 1998). A limitation to this approach is the need to quantify the yield of individual wells, usually by liquid scintillation counting. Although polyvinylidene fluoride (PVDF) filter assays have been used, the quantitative value of these is hindered by the high protein concentration of the cell-free extracts (>10 mg/ml) and the limited capacity of the PVDF membranes so that additional dilution steps are required before sample application.

NMR can give structural information in solution and this is already being used to characterise and screen protein libraries for folded states. This is most frequently achieved in structural genomics projects using the ^{1}H-^{15}N HSQC spectrum, which in principle, yields one resonance per non-proline residue and some readily identified signals due to side chains, the dispersion of these signals in both dimensions (and in particular the ^{1}H dimension) providing an indication of protein folding (Yee et al., 2002).

Simple one-dimensional proton spectra have also been widely used to assess protein folding, with information being obtained by the examination of specific spectral regions. The dispersion of the amide ^{1}H signals, the presence of high-field shifted methyl resonances (δ 1.0 to -1.0) or H_{α} resonances in the region characteristic of β-sheet structure (δ 5.0–6.0) all indicate that a protein is correctly folded (Rehm et al., 2002). While ^{15}N labelling is inexpensive, the resulting NMR spectra are relatively complex, especially for proteins of molecular weight >15 kDa.

Due to the range and dependence on the local environment of ^{19}F chemical shifts (which is greater than that of ^{1}H) and the intrinsic sensitivity of the fluorine nucleus (83% that of ^{1}H), it is realistic to expect that 1D-^{19}F spectra can also provide information on the extent of protein folding. This property has been exploited in peptide and small molecule HTP screening (Dalvit et al., 2003) and many fluorinated amino acids, e.g., 5 and 6-fluorotryptophan, are readily available. Proteins labelled with fluorinated tryptophan yield extremely simple ^{19}F spectra, with one resolved signal occurring for each tryptophan residue. The signals are very sensitive to environment, and thus to the folded state. If more than one conformation exists, separate sets of signals are often detected for each significant population (Gerig, 1994). By contrast, the resonances of an unfolded protein (in which the tryptophan residues all experience a similar environment), are essentially degenerate. There is considerable literature precedent for these observations using different fluorinated tryptophan analogues (Sykes et al., 1974; Li et al., 1989; Hoeltzli and Frieden, 1994). Since 1D spectra can be acquired so rapidly and using such simple methods, in some situations ^{19}F spectra of fluoro-tryptophan labelled proteins might constitute a viable alternative to the HSQC for assessing protein folding.

In this paper we demonstrate a simple and fast screening method for the optimization of cell-free expression of a test protein, human serum retinol-binding protein (RBP), with 100% incorporation of 6-fluorotryptophan, and identification of the correctly folded protein by ^{19}F NMR.

Materials and methods

Construction of pET14b RBP

The expression vector pET14b RBP was generated by releasing the retinol binding protein (RBP) coding region from pETRBP (Greene et al., 2001) with NdeI and BamHI, and recloning in pET14b.

Cell-free protein expression

Protein expression was carried out using the cell free protein synthesis system developed by Prof. Shigeyuki Yokoyama's group (RIKEN, Yokohama, Japan). The system consists of a coupled transcription-translation reaction from a suitable construct using *Escherichia coli* S30 cell extract, T7 RNA polymerase and low molecular weight

substrates. The S30 extract was prepared from a BL21 codon plus RIL strain as described elsewhere (Kigawa et al., 2004).

For batch-mode expression in 96-well plates the reaction mix (30 µl) consisted of 55 mM Hepes-KOH pH 7.5, 4% polyethylene glycol (PEG) 8000, 210 mM potassium glutamate, 1.8 mM DTT, 1.2 mM ATP, 0.8 mM each of CTP, GTP, UTP, 0.64 mM 3′,5′-cyclic AMP, 35 µg/ml folinic acid, 27.5 mM ammonium acetate, 80 mM creatine phosphate, 0.25 mg/ml creatine kinase, 175 µg/ml *Escherichia coli* total tRNA, 0.05% sodium azide, 10.7 mM magnesium acetate, 1 mM each amino acid with tryptophan replaced with 0.25–4 mM 6-fluoro-DL-tryptophan (Sigma), 0.5 mM methionine, 0.5 µl of ^{35}S labelled methionine (15 mCi/ml Amersham Bioscience), 0.27 µl T7 RNA polymerase (200 U/µl, Ambion), 7.2 µl S30 extract and 60–250 ng of pET14b RBP template DNA. The reactions were incubated within a polypropylene 96-well plate (Anachem) in a Dyad DNA Engine thermocycler (MJ Instruments).

To compare the effects of fluorinated tryptophan on protein expression, batch mode reactions were set up as above with 1 mM tryptophan, 5-fluoro-D-tryptophan, and 6-fluoro-DL-tryptophan, and 250 ng of plasmids encoding green fluorescent protein (GFP) (pIVEX GFP from Roche Diagnostics), RBP, 9th and 10th type III fibronectin repeats (9,10FIII) (pRSET FIII9-10 from Dr H. Mardon, Oxford) and hemolysin (HL) (pT7 HL from Prof H. Bayley, Oxford).

For large scale expression, 3 ml of the reaction mix without radioactive isotope was placed in a dialysis bag (Spectra/Por 2.1, 50 kDa MWCO) in 30 ml of external solution consisting of the same composition as the reaction mix except for the creatine kinase, the plasmid DNA, the T7 RNA polymerase, the S30 extract and also containing an additional 4.2 mM magnesium acetate. For fluorine labelling the tryptophan in the external and internal solutions was replaced with 1mM 6-fluoro-DL-tryptophan. The reaction was incubated at 30 °C at 160 rpm for 12 h.

Filter assay

To quantify the expression, 2 µl of ^{35}S labelled reaction mixture was applied to a Type GF/C glass fibre filter (Whatman) after centrifugation of the samples at 6,500g for 30 minutes in a Beckman JS-5.9 rotor. The proteins were precipitated and free amino acids including radioactive label removed with TCA (10%, w/v) in sodium pyrophosphate (1%, w/v) by washing for 10 min and then twice with TCA (5%, w/v) for 5 min. The washed filters were rinsed with methanol to facilitate drying at room temperature. The dried filters and polyacrylamide gel electrophoresis (PAGE) gels were exposed for 10 min to 20 by 25 cm general purpose phosphor screens which were subsequently read with a Storm 820 phosphorimager and the images processed with Image Quant software (Amersham Bioscience). Grid objects were created that matched the position of spots on the filter, and the volume reports calculated. The volume reports for a fixed time interval are directly proportional to the radioactivity and hence to the number of methionines present in the sample. This quantitative data was transferred to an Excel spreadsheet for further processing. Background and negative control samples were subtracted from all samples and the mean and standard deviations calculated for each triplicate. For comparing volume reports from different proteins, the values were normalized by dividing by the number of methionines in the protein constructs.

Protein purification and characterization

Protein samples were run on precast 10% Bis Tris polyacrylamide Novex gels in MES running buffer (Invitrogen). Radioactive samples were exposed to phosphor screens and processed as before, while non-radioactive samples were stained with Coomassie Brilliant Blue.

The large scale reaction mixture from the dialysis bag was buffer exchanged to 50 mM Tris, pH 7.5, 200 mM NaCl with a Centricon Ultrafree concentrator MWCO 10 kDa (Millipore) and applied to a HisTrap 1 ml column (Amersham Biosciences) equilibrated with the same buffer and the His-tagged protein eluted with an imidazole gradient (0–0.5 M) in the same buffer.

Purified protein was N-terminally sequenced on an Applied Biosystem 494A 'Procise' Sequencer and the molecular weight determined by electrospray ionization mass spectroscopy on a VG Platform II ESI-MS. Homogeneity was

determined and protein concentration estimated by capillary electrophoresis under non-reducing conditions on a 2100 Bioanalyzer using the Protein 50 Kit (Agilent Technologies).

^{19}F NMR

^{19}F NMR spectra were obtained at 564 MHz on a home-built spectrometer consisting of an Oxford Instruments magnet and a GE Omega console. Samples were prepared in 5 mM sodium phosphate at pH 7.0, with the 1 mM 6-fluoro-DL-tryptophan standard and recombinant refolded protein in a 5 mm Wilmad sample tube and the cell-free product in 290 µl in a Shigemi tube. About 5% D_2O was added to provide a lock signal. As an example of an unfolded protein spectrum the cell-free sample was diluted 1:1 with 8 M guanidine hydrochloride. Spectra were acquired at either 20 °C (refolded) or 25 °C (label standard and cell-free), standard spectral parameters being a 10,000 Hz spectral width, 4096 complex points, 90° pulse width, 1.0 s relaxation delay. No proton decoupling was applied during acquisition. Spectra were processed using FELIX 2.3 (Biosym Technologies Inc.), each FID was zero-filled to 8192 points and a line broadening of 20 Hz was applied. The spectra were indirectly referenced against TFA at 0 ppm (Maurer and Kalbitzer, 1996).

Results and discussion

While cell-free translation systems have been developed for increased protein yields, screening of the small scale expression, usually in 96-well format, has become a limiting factor in the process of optimizing conditions and constructs for protein expression. PAGE analysis allows the identification and quantification of protein bands but is costly, labour intensive and time consuming. The identification of proteins by fluorescent or antibody labelling on dot blots has been described (Beernink et al., 2003; Busso et al., 2003). Our experience with similar PVDF membrane processes indicates that although these approaches can identify the presence or absence of the target protein they are not quantitative enough to allow optimisation of conditions, before the scaling up of expression to produce samples suitable for biophysical analysis. The filter assay does not depend on the protein binding capacity of the filter medium but instead uses TCA to precipitate the entire protein sample within the inert filter matrix. The sensitivity and dynamic linearity of the phosphor screens allows the quantification of sample radioactivity comparable to liquid scintillation counting but with no dilution, additional liquid handling steps or requirement for scintillant. The filter assay usually takes less than an hour to complete and analyse, allowing the optimization and expression scale-up to be accomplished in the same day. The process can be automated using a pin-array plate replicator common to many laboratory robot systems.

Human RBP was chosen as the test protein for these experiments as a ^{19}F NMR spectrum of the 6-fluorotryptophan labelled and correctly folded protein was available. This protein sample was produced using the same techniques that generated an X-ray crystal structure (Greene et al., 2001). The protein has four tryptophan residues (positions 24, 67, 91 and 105), a molecular weight of 21.2 kDa and three disulphide bridges in the native β-barrel structure. The ^{19}F NMR spectrum of refolded RBP was obtained by introducing the fluorinated amino acid in the growth media for a bacterial strain expressing the gene. The protein was purified and refolded from inclusion bodies (Greene et al., 2001), and a ^{19}F NMR spectra acquired yielding four well dispersed fluorine peaks (Figure 3d). In addition to the lengthy processes involved, the labelling efficiency was approximately 90% with an uneven distribution over the four amino acid positions. This is consistent with other experiments using such classical labelling with reports of incorporation as low as 42% (Abbott et al., 2004; Senear et al., 2002). To allow purification from the cell-free reaction mixture the RBP construct was subcloned into pET14b adding 20 amino acids, including six histidines and a thrombin cleavage site, to the amino terminus and increasing the molecular weight to 23.4 kDa. This His-tag allowed an efficient, one step purification of the RBP.

The conditions of temperature (15, 20, 30 and 40 °C), DNA template concentration (2.0, 4.1, and 8.3 µg/ml), and 6-fluoro-DL-tryptophan (0.25, 0.5, 1, 2 and 4 mM) were explored simultaneously using cell-free expression in a microtitre plate

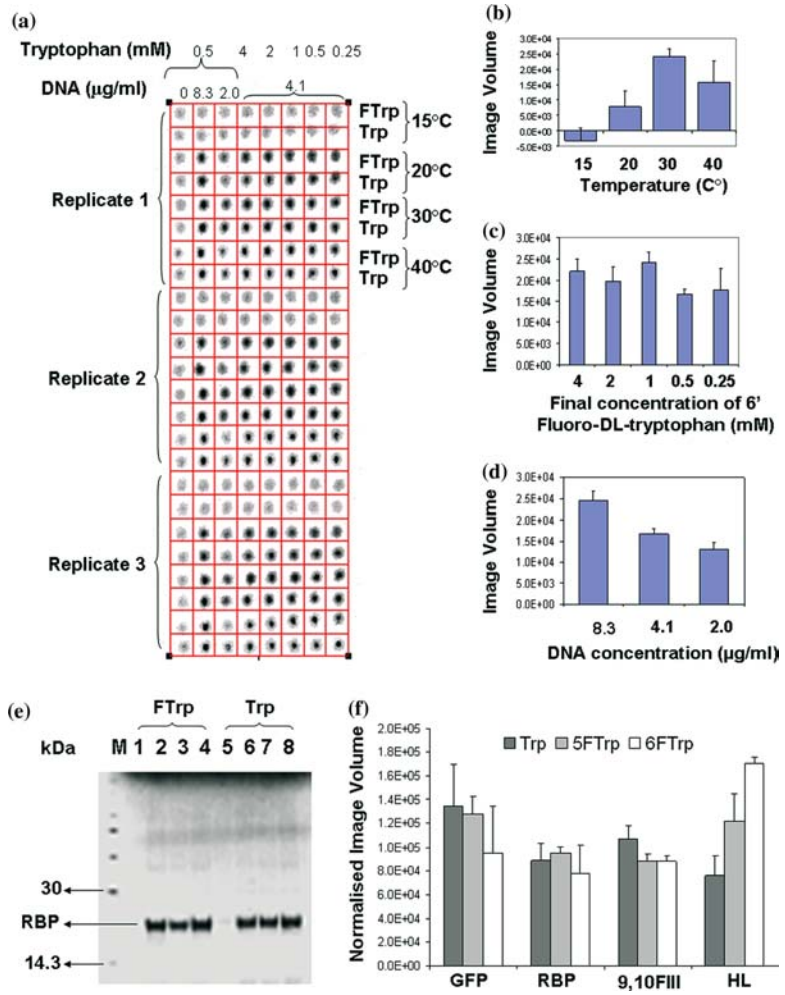

Figure 1. (a) Phosphor screen image of the filter assay for comparison of the effects of temperature, 6-fluoro-DL-tryptophan concentration and DNA concentration on ^{35}S labelled RBP cell-free expression levels. Plasmid DNA and tryptophan concentration were varied in columns while 6-fluoro-DL-tryptophan (FTrp), tryptophan (Trp) and temperature (15–40 °C) were varied as rows. Samples with no DNA (first column) were used as negative controls. (b), (c) and (d) Graphical representation of filter assay data from averaged ImageQuant volume reports plotted against temperature (at 1 mM 6-fluoro-DL-tryptophan, 4.1 µg/ml plasmid DNA), 6-fluoro-DL-tryptophan concentration (at 30 °C, 4.1 µg/ml plasmid DNA), and DNA concentration (at 0.5 mM 6-fluoro-DL-tryptophan, 30 °C) in the cell-free expression mix. Error bars represent standard deviation over the triplicate samples. (e) Phosphor screen image of 10% SDS PAGE for ^{35}S labelled cell-free samples showing only one product of approximately 24 kDa from the reactions. Lane M – ^{14}C methylated protein markers (Amersham, Biosciences); lanes 1 and 5 – negative controls (no plasmid DNA) of 2 mM tryptophan (Trp) and 1 mM 6-fluoro-DL-tryptophan (FTrp) respectively, lanes 2,3,and 4 – cell free translation performed with 4, 2 and 1 mM 6-fluoro-DL-tryptophan; lanes 6,7 and 8 – cell-free translation performed with 6, 4 and 2 mM of tryptophan respectively. All batch reactions were conducted at 30 °C and 8.3 µg/ml DNA where required. (f) Graphical representation of filter assay data from averaged ImageQuant volume reports for a comparison of relative protein yields from batch reactions of GFP, RBP, 9,10FnIII and HL with 1 mM tryptophan (Trp), 5-fluoro-L-tryptophan (5FTrp) and 6-fluoro-DL-tryptophan (6FTrp) at 30 °C, 8.3 µg/ml plasmid DNA. The signals from triplicate samples were normalized by dividing by the number of methionine residues in each protein construct. Error bars represent standard deviation over the triplicate samples.

using the filter assay for analysis (Figures 1a–d). The maximum protein yield for each condition was selected (30 °C, 8.3 µg/ml DNA and 1 mM 6-fluoro-DL-tryptophan) and a PAGE was run with the samples and negative controls to ensure that there was minimal background and no protein degradation, confirming the results of the filter assay (Figure 1e). A comparison of protein

expression with tryptophan, 5-fluoro-L-tryptophan, and 6-fluoro-DL-tryptophan for GFP, RBP, 9,10FnIII, and HL (Figure 1f) indicates that yields are protein dependant with no overall trends. While the expression of RBP and 9,10FnIII show little difference between the forms of tryptophan, GFP exhibits reduced expression with 5- and 6-flurotryptophan, but hemolysin shows an increase with the fluorinated amino acids. This suggests that expression conditions should be optimized for each protein and fluorinated amino acid combination, rather than relying on a set of standard conditions for all proteins.

The solubility of the RBP (approximately 90% of the total RBP expressed) was not significantly affected by the various conditions. A large scale expression was set up with these conditions and yielded after purification 0.3 mg of RBP per 3 ml reaction (Figure 2a). The RBP was purified through the His-tag (Figure 2a) and the homogeneity of the sample was confirmed by capillary electrophoresis under non-reducing conditions (Bousse et al., 2001) (Figure 2b). The purified protein gave a single band or peak under both reducing and non reducing conditions suggesting that any disulphides formed were intramolecular.

Electrospray ionization mass spectrometry analysis, interpreted with the deconvolution algorithm MaxEnt, determined a molecular weight of 23,460 Da for the purified protein (Figure 2c) consistent with a protein with 100% substitution of the four tryptophans with fluorotryptophan, an amino terminal formylmethionine and three disulphide bonds (predicted molecular weight 23,460 Da). These results suggest that the RBP has undergone folding with the formation of disulphide bonds but does not confirm a single folded species. Amino terminal sequencing of the full length protein was unsuccessful probably due to the presence of a formyl group blocking the amino terminus but a small fraction of the sample (approximately 5% of the total protein) gave the amino sequence SGLVPRGSHMERD that agrees with residues 12–24 of the pET14b construct. This minor component of the sample appears to have lost eleven amino terminal residues after purification due to contaminating proteases.

For early characterization of protein sample suitability for structural genomics, rather than NMR assignment, ^{19}F NMR has advantages over ^{15}N HSQC screening since the range and sensitivity of fluorine chemical shifts are much higher for fluorine than hydrogen. The fluorine nucleus produces a strong NMR signal (83% that of ^{1}H) against a background devoid of signals from endogenous fluorine and the ^{19}F chemical shifts of fluorinated tryptophan are essentially degenerate in the denatured state, but once tertiary structure is formed there are large shifts both up and down field due to the local differences in the individual environments of the fluorinated residues. Because of these conditions, a 50 μM sample of the cell-free expressed RBP was sufficient to obtain a useful ^{19}F spectrum which can be compared against that of the free label, 6-fluoro-DL-tryptophan, to identify chemical shifts indicative of folding. While a 6-fluorotryptophan sample (**FTrp**, Figure 3a) and an unfolded RBP sample (**U**, Figure 3b) gave spectra consisting of overlapping peaks centred at δ −43.77 and −43.6 respectively, both the cell-free sample and the refolded sample exhibit nearly identical spectra with signals at δ −40.3 (I), −43.2 (II), −44.1 (III), −44.6 (IV) cell-free (Figure 3c), δ −40.3 (I), −43.2 (II), −44.1 (III), −44.8 (IV) refolded (Figure 3d), indicating that the proteins have the same tertiary structure. The sharp signal at δ −41.6 (*) in the cell-free spectrum is due to fluoride ion contamination of the buffer solutions, and has a chemical shift that agrees with previous reports (Martinez et al., 1996). With the conditions used for the denatured sample (Figure 3b) this fluoride signal moves downfield out of the range used for the figure. The incorporation of the fluorotryptophan thus allows unambiguous identification of unfolded versus folded protein (Figures 3b and c) and, in addition, confirms that the optimized cell-free expression has expressed correctly folded, disulphide bonded RBP without the requirement of any refolding steps (Figures 3c and d). The spontaneous formation of disulphides in the cell-free expression of RBP was unexpected but has previously been reported for other cell-free expression systems (Ryabova et al., 1997; Kim and Swartz, 2004).

The high efficiency of 6-fluorotryptophan incorporation is comparable to that observed with selenomethionine and isotope-labelling

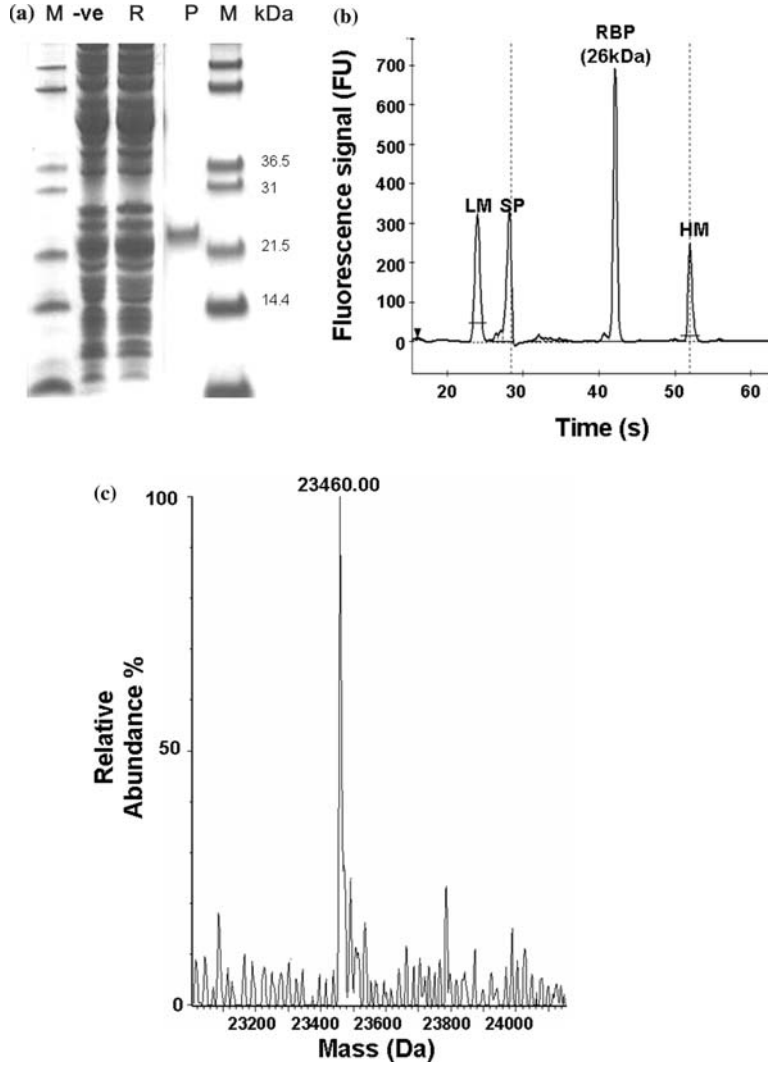

Figure 2. (a) 10% SDS PAGE analysis under reducing conditions of cell-free expression reaction mix after 12 h incubation without pET14b RBP template (−ve), with pET14b RBP template (R) , 6-fluorotryptophan-RBP after HisTrap column purification (P), and MW markers (M). The proteins were visualized with Coomassie Brilliant Blue. (b) Capillary electrophoresis of 6-fluorotryptophan-RBP under non-reducing conditions. The system peak (SP) and low (LM) and high (HM) MW markers are common to all runs with the sample running between SP and HM. The protein size is estimated by comparison to protein standards (6–53 kDa) and the sample concentration by comparison of peak area with that of HM. The 11% discrepancy between the predicted MW (23.46 kDa) and that calculated from mobility (26.0 kDa) is within the observed variation for samples in this system. (c) The deconvoluted ESI-mass spectrum of 6-fluorotryptophan-RBP. The calculated MW (23, 460 Da) of the major species exactly agrees with an amino-terminal formylmethionine, disulphide bonded, 100% fluorotryptophan-labelled RBP as expressed from pET14b RBP.

reported for cell-free systems (Kigawa et al., 2001). Although 6-fluorotryptophan was used in this report because of the availability of a suitable NMR spectrum for comparison, other commercially available fluorinated amino acids, e.g., 4-fluoro-L-phenylalanine, could be used.

These experiments have demonstrated that a ^{19}F spectrum of a 6-fluorotryptophan labelled sample can identify correctly folded protein. The NMR methodology involves relatively simple experiments and processing, which can (depending upon the quantity of protein produced) either allow extremely rapid assessment of the folded/unfolded state or enable the study of small amounts of material. The ^{19}F spectrum is also simple to interpret since only resonances

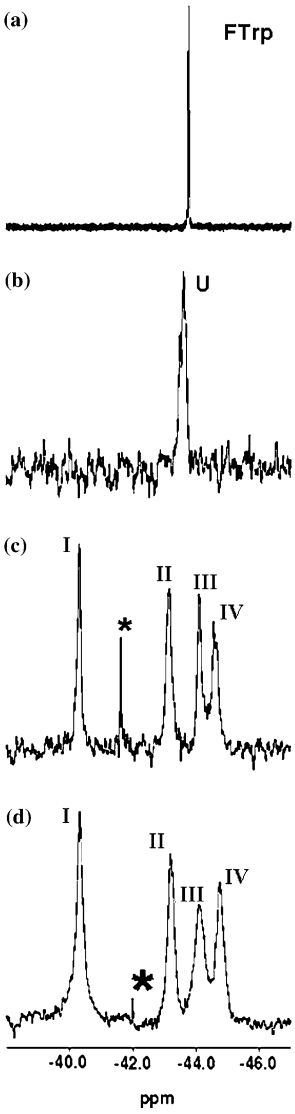

Figure 3. Comparison of ^{19}F NMR spectra of 6-fluoro-DL-tryptophan 1 mM at 25 °C (a), unfolded 6-fluorotryptophan-RBP 25 μM at 25 °C (b), purified cell-free expressed 6-fluorotryptophan-RBP 50 μM at 25 °C (c) and refolded 6-fluorotryptophan-RBP at 20 °C (d) in 5 mM sodium phosphate pH 7.0. The spectra are referenced against TFA at 0 ppm. A total of 32, 58, 000 and 65, 000 scans were acquired for a, b and c respectively.

corresponding to the tryptophan residues are observed, which allows the total number of signals to be readily determined and should assist in the identification of multiple conformations and allow mixtures of folded and unfolded protein to be distinguished from homogeneous, partially folded samples. This contrasts with the 1D ^{1}H spectrum that contains thousands of resonances from all residues in the protein and can be further complicated by residual solvent signals (and in some cases, signals from the buffer).

^{19}F NMR might also possess certain advantages over ^{1}H-^{15}N HSQC screening for the early identification of suitable protein samples for structural genomics studies, rather than full NMR assignment. While ^{15}N labelling is inexpensive, the resulting NMR spectra are relatively complex, especially for proteins of molecular weight >15 kDa. The acquisition of suitable spectra also becomes more challenging as the size of the protein increases, requiring the implementation of TROSY methods above 20 kDa. Furthermore, suppression of the residual water signal can be problematic in some cases and at pH above 7.0, fast exchange between the amide protons and the bulk solvent can significantly reduce the intensity of the HSQC resonances.

In contrast, acquiring and processing a 1D spectrum in the ^{19}F approach is extremely simple, the 6-fluorotryptophan resonances are unaffected by solvent exchange allowing the use of physiological pH and the background is devoid of signals from endogenous fluorine. 1D spectra can also be acquired substantially faster than a 2D HSQC, thus either reducing the experiment time or allowing the use of small quantities of material. Because of these conditions, a 50 μM sample of the cell-free expressed RBP was sufficient to obtain a useful ^{19}F spectrum. Even if the sample protein contains only one tryptophan residue, comparison of the ^{19}F spectrum to that of the free label will allow chemical shifts due to folding to be identified. The technique can therefore be applied to samples for which no folded spectra are available.

In combination with the improvements in batch-mode cell-free expression levels (Sawasaki et al., 2002), the ability to obtain NMR spectra from labelled proteins in cell-free expression reactions without the need to purify the protein (Guignard et al., 2002) and the increasing sensitivity of cryoprobes (Styles et al., 1989) it should be feasible to identify folded proteins by NMR from cell-free samples expressed at the microlitre scale in microtitre plates. Such an approach would go far to removing a major bottle-neck in structural studies.

Acknowledgements

We gratefully acknowledge the support from the following organisations – National Overseas Fellowship, Govt. of India (MN), NSF International Research Fellows Award, Grant #0000597 and BBSRC Grant 43/B17889 (LHG), and the Wellcome Trust (SAC, IDC, DS). We thank Prof. S. Yokoyama and his group for the cell-free protein synthesis protocols and assistance, Tony Willis for N terminal sequencing and Nick Soffe for NMR instrumentation.

References

Abbott, G.L., Blouse, G.E., Perron, M.J. et al. (2004) *Biochemistry*, **43**, 1507–1519.

Beernink, P.T., Segelke, B.W., Coleman, M.A. (2003) In *Cell-Free Protein Expression*, Swartz, J.R. (Ed.), Springer-Verlag, Berlin, pp. 101–107.

Bousse, L., Mouradian, S., Minalla, A. et al. (2001) *Anal. Chem.*, **73**, 1207–1212.

Busso, D., Kim R. and Kim S.-H. (2003) *J. Biochem. Biophys. Meth.*, **55**, 233–240.

Christendat, D., Yee, A., Dharamsi, A. et al. (2000) *Nat. Struct. Biol.*, **7**, 903–909.

Dalvit, C., Fagerness, P.E., Hadden, D.T.A. et al. (2003) *J. Am. Chem. Soc.*, **125**, 7696–7703.

Gerig, J.T. (1994) *Prog. Nucl. Magn. Reson. Spectrosc.*, **26**, 293–370.

Greene, L.H., Chrysina, E.D., Irons, L.I. et al. (2001) *Protein Sci.*, **10**, 2301–2316.

Guignard, L., Ozawa, K., Pursglove, S.E. et al. (2002) *FEBS Lett.*, **524**, 159–162.

Hoeltzli, S.D. and Frieden, C. (1994) *Biochemistry* **33**, 5502–5509.

Kigawa, T. and Yokoyama, S. (1991) *J. Biochem. (Tokyo)* **110**, 166–168.

Kigawa, T., Muto, Y. and Yokoyama, S. (1995) *J. Biomol. NMR*, **6**, 129–134.

Kigawa, T., Yabuki, T. and Matsuda, N., et al. (2004) *J. Struct. Funct. Genom.*, **5**, 63–68.

Kigawa, T., Yabuki, T., Yoshida, Y. et al. (1999) *FEBS Lett.*, **442**, 15–19.

Kigawa, T., Yamaguchi-Nunokawa, E. and Kodama, K., et al. (2001) *J Struct. Funct. Genom.*, **2**, 27–33.

Kim, D.M. and Swartz, J.R. (2004) *Biotechnol. Bioeng.*, **85**, 122–129.

Li, E., Qian, S.J., Nader, L. et al. (1989) *J. Biol. Chem.*, **264**, 17041–17048.

Martinez, E.J., Girardet, J.-C. and Morat, C. (1996) *Inorg. Chem.*, **35**, 706–710.

Maurer, T. and Kalbitzer, H.R. (1996) *J. Magn. Reson. B*, **113**, 177–178.

Rehm, T., Huber, R. and Holak, T.A. (2002) *Structure*, **10**, 1613–1618.

Ryabova, L.A., Desplancq, D., Spirin, A.S. et al. (1997) *Nat. Biotechnol.*, **15**, 79–84.

Sawasaki, T., Hasegawa, Y., Tsuchimochi, M. et al. (2002) *FEBS Lett.*, **514**, 102–105.

Senear, D.F., Mendelson, R.A., Stone, D.B. et al. (2002) *Anal. Biochem.*, **300**, 77–86.

Staunton, D., Owen, J. and Campbell, I.D. (2003) *Acc. Chem. Res.*, **36**, 207–214.

Styles, P., Soffe, N. and Scott, C. (1989) *J. Magn. Reson.*, **84**, 376–378.

Sykes, B.D., Weingarten, H.I. and Schlesinger, M.J. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 469–473.

Yabuki, T., Kigawa, T., Dohmae, N. et al. (1998) *J. Biomol. NMR*, **11**, 295–306.

Yee, A., Chang, X.Q., Pineda-Lucena, A. et al. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 1825–1830.

Yokoyama, S. (2003) *Curr. Opin. Chem. Biol.*, **7**, 39–43.