# NMR analysis of in vitro-synthesized proteins without purification: a high-throughput approach

Laurent Guignard[a,b], Kiyoshi Ozawa[a,b], Sharon E. Pursglove[a,b], Gottfried Otting[a,b], Nicholas E. Dixon[a,*]

[a] *Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia*
[b] *Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden*

Received 4 June 2002; accepted 18 June 2002

First published online 8 July 2002

Edited by Horst Feldmann

**Abstract A cell-free protein expression system was established that provides protein samples of adequate concentration and purity for direct NMR analysis. The *Escherichia coli* peptidyl–prolyl *cis–trans* isomerase PpiB was expressed in this system with dual amino acid-selective isotope labeling to identify the NMR signal from the active site-residue Arg87. Addition of the substrate succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide selectively shifted its $^{15}$N-HSQC cross peak, confirming binding to the active site. As cell-free protein expression provides high yields of protein per unit mass of labeled amino acid and sample handling is minimal, this strategy presents an exceptionally inexpensive and rapid approach to protein analysis.** 

*Key words:* In vitro protein expression; Cell-free; High throughput; PpiB; Prolyl *cis–trans* isomerase

## 1. Introduction

NMR spectroscopy is an important tool for the identification of protein–ligand interactions. Often, the $^{15}$N-HSQC spectrum of a sample of $^{15}$N labeled protein is used to monitor ligand binding. Since every backbone amide group is represented by a cross peak in the $^{15}$N-HSQC spectrum, chemical shift changes or the disappearance of signals observed upon addition of the ligand provide site-specific information about ligand binding [1]. With prior knowledge about potential binding sites, the $^{15}$N-HSQC cross peaks of strategically chosen amino acids can be identified by a dual amino acid-selective labeling technique to avoid assignment of the entire $^{15}$N-HSQC spectrum [2–4]. A major drawback of the selective labeling strategy for proteins expressed in vivo is the time-consuming purification of every selectively labeled sample and the higher cost of labeled amino acids versus the ammonium salts used for uniform labeling with $^{15}$N. Various strategies to increase efficiency of such experiments have focused on recording $^{15}$N-HSQC spectra of partially purified proteins or even in situ in whole *Escherichia coli* cells [5–10]. Particularly clean spectra have been obtained by the use of auxotrophic bacterial strains for expression [11]. However, in-cell NMR results in broad signals due to high intracellular viscosity, while cell lysis exposes the protein to temperature and chemical stress, including exposure to cell debris and proteolytic enzymes.

An attractive alternative is presented by use of a cell-free in vitro protein expression system, which can produce high yields of protein per input mass of labeled amino acid without the need for use of auxotrophic bacterial strains [4,12–14]. Here we show that cell-free protein expression produces selectively labeled proteins in adequate yield and of high purity so that, when combined with the sensitivity of a cryoprobe, $^{15}$N-HSQC spectra can be recorded directly without further purification or concentration. This opens the door to high-throughput ligand-binding studies by NMR spectroscopy, where the number of sample handling steps is drastically minimized and parallel sample preparation is readily amenable to automation.

## 2. Materials and methods

Two *E. coli* proteins were expressed, cytoplasmic peptidyl–prolyl *cis–trans* isomerase (PpiB) and aspartyl-tRNA synthetase (AspRS), using modifications of the cell-free expression procedure of Kigawa et al. [14]. Supercoiled plasmids pKE874 [15] and pPL830 (P.E. Lilley and N.E.D., unpublished) were used as DNA templates for in vitro expression of PpiB and AspRS, respectively. These plasmids contain the *E. coli ppiB* and *aspS* genes inserted so that they are transcribed from tandem bacteriophage λ $p_R$ and $p_L$ promoters [16]. Unlabeled PpiB was purified as described [15]. Its concentration was determined spectrophotometrically using a molecular weight of 18 154 and an $\varepsilon_{280}$ value of 9 530 $M^{-1}$ $cm^{-1}$ [17]. *E. coli* RNA polymerase was purified from cells of strain C600 as described by Burgess et al. [18] and separated into holoenzyme and core fractions as described by Gonzalez et al. [19].

The S-30 extract used for the cell-free protein synthesis was prepared from *E. coli* strain A19 [20] (*metB rna*; CGSC No. 5997), as described [21]. The S-135 extract and a crude ribosomal fraction were prepared from the S-30 extract according to Patnaik and Swartz [22], and Kudlicki et al. [23], respectively. [U-$^{15}$N]L-arginine and [1-$^{13}$C]L-alanine were gifts of Cambridge Isotope Laboratories (Andover, MA, USA).

Inner-chamber reaction mixtures (0.3 ml each) contained HEPES–KOH (55 mM, pH 7.5), dithiothreitol (1.7 mM), ATP (1.2 mM), CTP, GTP and UTP (0.8 mM each), 3′,5′-cyclic AMP (0.64 mM), folinic acid (68 μM), ammonium acetate (27.5 mM), potassium glutamate (208 mM), creatine phosphate (80 mM), creatine kinase (Roche; 250 μg/ml), twenty L-amino acids (1 mM each), magnesium acetate (15 mM), *E. coli* total tRNA (Roche; 175 μg/ml), $NaN_3$ (0.05%, w/v), S-135 extract (48 μl), ribosome fraction (24 μl), supercoiled plasmid DNA (16 μg/ml for PpiB; 14 μg/ml for AspRS), RNase inhibitor (Promega; 150 U), and RNA polymerase holoenzyme (0.155 mg/ml). When used, PEG-8000 was added to 2% (w/v). The inner-chamber reaction mixtures were dialyzed in Spectrapor 2 tubing (12–14 kDa

*Corresponding author. Fax: (61)-2-6125 0750.
*E-mail address:* dixon@rsc.anu.edu.au (N.E. Dixon).

0014-5793/02/$22.00 
PII: S0014-5793(02)03048-X

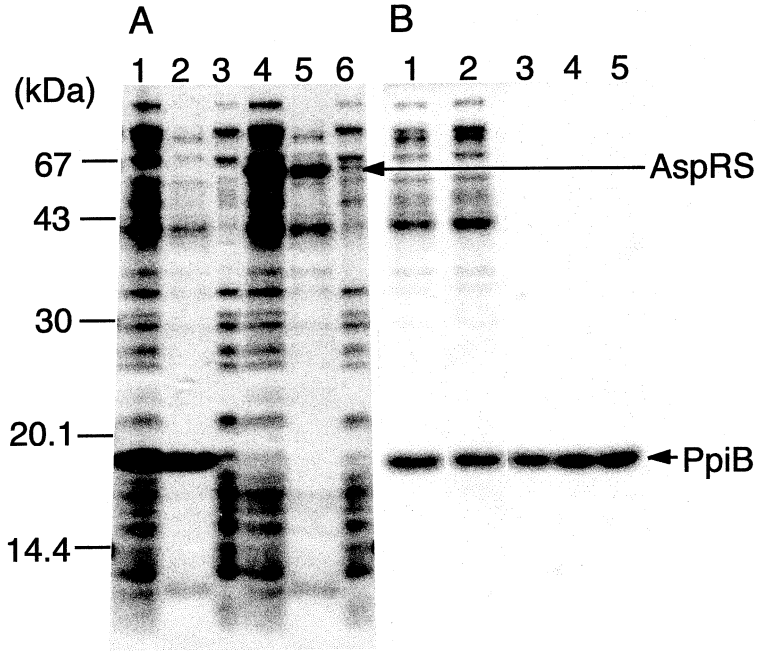


Fig. 1. In vitro expression of PpiB (18 kDa) and aspartyl-tRNA synthetase (AspRS, 66 kDa). Reactions were carried out for 12 h using a dialysis system, and products were separated by SDS–polyacrylamide (15%) gel electrophoresis. A: Expressed proteins are soluble and reaction mixtures can be purified by ultracentrifugation. Reaction mixtures contained 2% (w/v) PEG-8000. Lanes 1–3: PpiB; lanes 4–6: AspRS. Lanes 1 and 4: crude inner-chamber reaction mixtures; lanes 2 and 5, supernatants after ultracentrifugation (PpiB and AspRS give the strongest bands); lanes 3 and 6: pellets after ultracentrifugation (PpiB and AspRS are virtually absent). B: Expression of PpiB is unaffected by omission of PEG-8000. Supernatants after ultracentrifugation of crude inner-chamber mixtures from reactions carried out in the presence (lane 1) or absence (lane 2) of PEG-8000 (2% w/v). These were the samples of $^{13}$C-Ala/$^{15}$N-Arg labeled PpiB used for NMR experiments (Fig. 2). Lanes 3, 4 and 5 from the same gel contained amounts of purified PpiB corresponding to final concentrations in the reaction mixtures of 0.8, 1.0 and 1.2 mg/ml, respectively. Mobilities of molecular weight markers were as indicated. The gels were stained with Coomassie Blue.

cutoff; Spectrum Laboratories), for 12 h at 37°C with shaking at 180 rpm against outer-chamber solution (6 ml; changed after 3 and 6 h) contained in a test tube. The outer-chamber solution had the same composition as the inner-chamber reaction mixture, except that S-135 extract, ribosomal fraction, tRNA, plasmid DNA, RNA polymerase, creatine kinase and RNase inhibitor were omitted, and the concentration of magnesium acetate was raised to 19.3 mM. Prior to analysis by sodium dodecyl sulfate (SDS)–polyacrylamide (15%) gel electrophoresis, proteins in samples were precipitated with five volumes of ice-cold acetone, collected by centrifugation, resuspended in a gel loading mixture containing 2% (w/v) SDS, and heated for 2 min at 90°C.

For preparation of samples of $^{15}$N-Arg/$^{13}$C-Ala labeled PpiB, the labeled amino acids (1 mM each) replaced unlabeled L-Ala and L-Arg in both mixtures. NMR measurements were carried out after combining two identical 0.3 ml reaction mixtures, followed by their ultracentrifugation (100 000× $g$, 4 h) and addition of 10% $D_2O$ to provide a lock signal. All NMR spectra were recorded at 25°C, pH 7, on a Bruker DMX-600 NMR spectrometer equipped with a triple-resonance ($^1$H, $^{15}$N, $^{13}$C) cryoprobe with single-axis pulsed field gradients, using a conventional 5 mm sample tube.

## 3. Results and discussion

The reaction conditions were similar to those established by Kigawa et al. [14] with a few notable exceptions, including transcription driven by *E. coli* RNA polymerase from tandem bacteriophage λ $p_R$ and $p_L$ promoters in the supercoiled plasmid template DNA, rather than by phage T7 RNA polymerase from a T7 promoter. Furthermore, the extract was supplemented by a crude ribosomal fraction, rather than being preconcentrated, and additional purified *E. coli* RNA polymerase holoenzyme was added to boost expression yields. Finally, our system allowed the omission of polyethylene glycol (see below).

For two proteins expressed in this way, PpiB and AspRS, straightforward ultracentrifugation provided efficient purification, rendering both PpiB and AspRS as the predominant proteins in solution (Fig. 1A). Expression yields of $>1$ mg of PpiB per ml of reaction volume were achieved (Fig. 1B).

A solution of $^{15}$N-Arg/$^{13}$C-Ala labeled PpiB thus prepared was used to identify the cross peak of Arg87 by the $^1J(^{15}N,^{13}C)$ coupling between Arg87 and Ala86. PpiB contains five arginyl residues, of which only Arg87 is preceded by an alanyl residue. The $^{15}$N-HSQC spectrum in Fig. 2A shows the cross peaks of all five arginyl residues in $^{15}$N-Arg/$^{13}$C-Ala labeled PpiB. Recording of the first $^{15}$N–$^1$H plane of an HNCO experiment [24] resulted in a $^{15}$N-HSQC-type spectrum, containing only the cross peak of Arg87 (Fig. 2B). Addition of the PpiB substrate succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (suc-AAPF-*p*NA, purchased from Bachem, Switzerland) to 5 mM shifted and broadened the cross peak of Arg87, indicating a high population of the enzyme–substrate complex (Fig. 2C). This is as predicted from the $K_M$ values of ca. 1 mM measured for the related proteins, bovine and human cyclophilin, with the same substrate [25]. (To the best of our knowledge, $K_M$ values have not been measured for PpiB.) In contrast, the presence of 6 mM succinyl-Ala-Pro-Ala did not change the $^{15}$N-HSQC spectrum (data not shown), in spite of the complex of PpiB with succinyl-Ala-Pro-Ala-*p*NA being sufficiently stable for crystallization [26]. This suggests that the

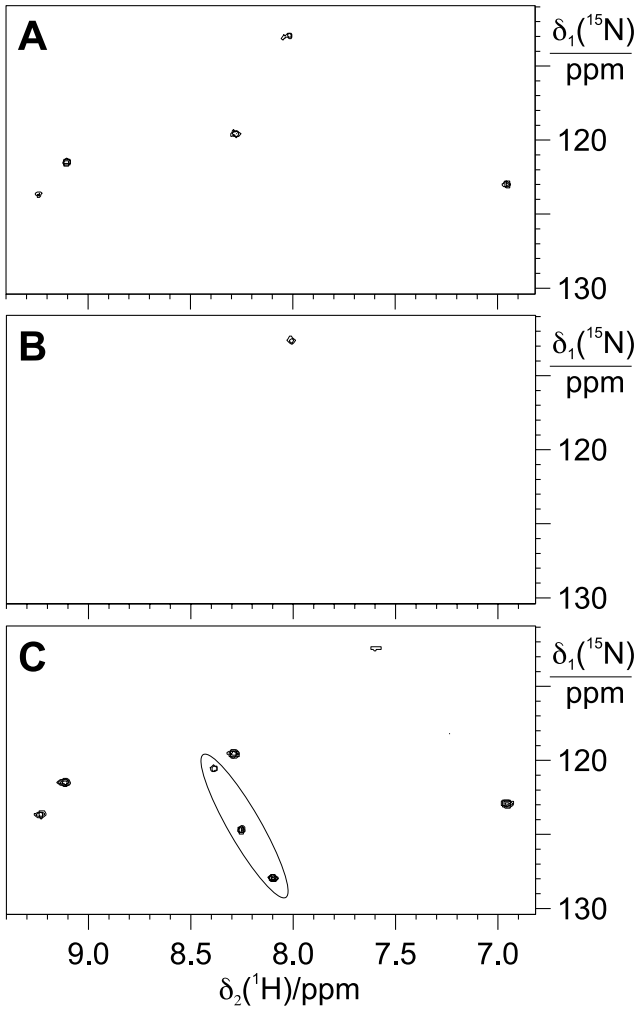


Fig. 2. $^{15}$N-HSQC NMR spectra of PpiB labeled with $^{15}$N-Arg and $^{13}$C-Ala. A: $^{15}$N-HSQC spectrum recorded in 2.7 h after ultracentrifugation and addition of $D_2O$ to 10% (v/v). B: Same as A, but with selection of $^{13}$C-bound $^{15}$N nuclei by an HNCO experiment [24], selecting the signal of Arg87. Total recording time of 4.4 h. C: Same as A, but after addition of 1.5 mg of the substrate succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide. The three new signals (circled) are from amide groups of the substrate at natural isotopic abundance. The total recording time was 12.8 h.

aromatic groups in this substrate, as well as in suc-AAPF-*p*NA, significantly stabilize the respective complexes with PpiB.

The high quality of the $^{15}$N-HSQC spectra is remarkable, as a conventional one-dimensional $^{1}$H NMR spectrum of the reaction mixture was completely dominated by the signals from buffer and amino acid components. Clearly, only the product protein was efficiently enriched with $^{15}$N-Arg and $^{13}$C-Ala. As expected, no cross peak was observed for any excess $^{15}$N-amino acid, since the signals from α-amino groups are broadened beyond detection by exchange with water protons. However, the *E. coli* extracts do contain enzymes that transform a fraction of the amino acids into low-molecular-weight metabolites, in which their α-amino groups are derivatized to amides. These compounds give rise to prominent signals in the $^{15}$N-HSQC spectra and were also observed in NMR spectra recorded of whole *E. coli* cells [10]. For example, the spectra in Fig. 2A,C contained a cross peak of narrow linewidth at 131.7 ppm ($^{15}$N)/7.87 ppm ($^{1}$H) (not shown). Although ultrafiltration readily separated the corresponding compound from PpiB, no ultrafiltration was carried out before recording the spectra shown in Fig. 2.

The cell-free expression system reported by Kigawa et al. [14] includes 4% PEG-8000 to achieve a concentration (excluded volume) effect. In the case of PpiB expressed with our cell-free system, however, very similar protein yields were obtained with and without PEG (Fig. 1B). The increased viscosity caused by 2% PEG resulted in less than 5 Hz line-broadening in the $^{1}$H dimension of the $^{15}$N-HSQC spectra.

## 4. Conclusions

Cell-free protein expression combined with the sensitivity of a cryoprobe enables an inexpensive high-throughput strategy for protein analysis: selectively labeled proteins can be expressed in 0.5 ml of reaction medium using small quantities of labeled amino acids (6.5 mg of $^{15}$N-Arg and 2.8 mg of $^{13}$C-Ala in the present case) and the samples can be analyzed by NMR and studied for ligand binding at the concentration of the reaction medium without any chromatographic purification and with minimal sample handling. All steps from expression to the ready NMR spectra can be completed in less than 24 h. Except for the broadened signal of Arg87, all cross peaks of Fig. 2C were already observable in a spectrum recorded in 2 h. Many samples could easily be prepared in parallel. Analysis of several selectively labeled samples is particularly attractive for large proteins, where uniform labeling would result in severe signal overlap and MUSIC-type residue-selective pulse sequences would fail due to sensitivity problems [27,28]. Notably, there is already sufficient signal overlap in the $^{15}$N-HSQC spectrum of uniformly labeled PpiB that the resonance of Arg87 is incompletely resolved, and it would not be trivial to attribute spectral changes induced by the substrate to specific residues [29].

*Acknowledgements:* We thank Cambridge Isotope Laboratories for gifts of $^{13}$C-Ala and $^{15}$N-Arg, Mark Mulcair for a sample of pure PpiB, the Wallenberg Foundation and the Wallenberg Functional Genomics Initiative for a grant to purchase the cryoprobe, and the EU (project QLCK3-CT-1999-00099), Fernström Foundation, Swedish Research Council and STINT for financial support.

## References

[1] Shuker, S.B., Hajduk, P.J., Meadows, R.P. and Fesik, S.W. (1996) Science 274, 1531–1534.
[2] Burk, S.C., Papastavros, M.Z., McCormick, F. and Redfield, A.G. (1989) Proc. Natl. Acad. Sci. USA 86, 817–820.
[3] Weigelt, J., Van Dongen, M., Uppenberg, J., Schultz, J. and Wikström, M.J. (2002) J. Am. Chem. Soc. 124, 2446–2447.
[4] Yabuki, T., Kigawa, T., Dohmae, N., Takio, K., Terada, T., Ito, Y., Laue, E.D., Cooper, J.A., Kainosho, M. and Yokoyama, S. (1998) J. Biomol. NMR 11, 295–306.
[5] Gronenborn, A.M. and Clore, G.M. (1996) Protein Sci. 5, 174–177.
[6] Almeida, F.C.L., Amorim, G.C., Moreau, V.H., Sousa, V.O., Creazola, A.T., Américo, T.A., Pais, A.P.N., Leite, A., Netto, L.E.S., Giordano, R.J. and Valente, A.P. (2001) J. Magn. Reson. 148, 142–146.
[7] Serber, Z., Keatinge-Clay, A.T., Ledwidge, R., Kelly, A.E., Miller, S.M. and Dötsch, V. (2001) J. Am. Chem. Soc. 123, 2446–2447.
[8] Serber, Z. and Dötsch, V. (2001) Biochemistry 40, 14317–14323.

[9] Serber, Z., Ledwidge, R., Miller, S.M. and Dötsch, V. (2001) J. Am. Chem. Soc. 123, 8895–8901.
[10] Lee, K.M., Androphy, E.J. and Baleja, J.D. (1995) J. Biomol. NMR 5, 93–96.
[11] Ou, H.D., Lai, H.C., Serber, Z. and Dötsch, V. (2001) J. Biomol. NMR 21, 269–273.
[12] Kigawa, T., Muto, Y. and Yokoyama, S. (1995) J. Biomol. NMR 6, 129–134.
[13] Kim, D.M., Kigawa, T., Choi, C.Y. and Yokoyama, S. (1996) Eur. J. Biochem. 239, 881–886.
[14] Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. and Yokoyama, S. (1999) FEBS Lett. 442, 15–19.
[15] Edwards, K.J., Ollis, D.L. and Dixon, N.E. (1997) J. Mol. Biol. 271, 258–265.
[16] Love, C.A., Lilley, P.E. and Dixon, N.E. (1996) Gene 176, 49–53.
[17] Gill, S.C. and von Hippel, P.H. (1989) Anal. Biochem. 182, 319–326.
[18] Burgess, R.R. and Jendrisak, J.J. (1975) Biochemistry 14, 4634–4638.
[19] Gonzalez, N., Wiggs, J. and Chamberlin, M.J. (1977) Arch. Biochem. Biophys. 182, 404–408.
[20] Clark, A.J. (1963) Genetics 48, 105–120.
[21] Pratt, J.M. (1984) in: Transcription and Translation (Hames, B.D. and Higgins, S.J., Eds.), pp. 179–209, IRL Press, Oxford.
[22] Patnaik, R. and Swartz, J.R. (1998) Biotechniques 24, 862–868.
[23] Kudlicki, W., Kramer, G. and Hardesty, B. (1992) Anal. Biochem. 206, 389–393.
[24] Grzesiek, S. and Bax, A. (1992) J. Magn. Reson. 96, 432–440.
[25] Kofron, J.L., Kuzmic, P., Kishore, V., Cõlon-Bonilla, E. and Rich, D.H. (1991) Biochemistry 30, 6127–6134.
[26] Konno, M., Ito, M., Hayano, T. and Takahashi, N. (1996) J. Mol. Biol. 256, 897–908.
[27] Schmieder, P., Leidert, M., Kelly, M. and Oschkinat, H. (1998) J. Magn. Reson. 131, 199–202.
[28] Schubert, M., Oschkinat, H. and Schmieder, P. (2001) J. Biomol. NMR 20, 379–384.
[29] Kariya, E., Ohki, S., Hayano, T. and Kainosho, M. (2000) J. Biomol. NMR 18, 75–76.