Protein and Peptide Letters, Vol. 10, No. 2, pp. 175-181, 2003
Bentham Science Publishers Ltd.
0929-8665/03 $ 41.00 + .00

AN EFFICIENT FUSION EXPRESSION SYSTEM FOR PROTEIN AND PEPTIDE OVEREXPRESSION IN *ESCHERICHIA COLI* AND NMR SAMPLE PREPARATION

Yuan Cheng, Dongsheng Liu, Yanming Feng, Guozhong Jing*

National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, China
*Author for correspondence (Fax: +8610-64872026; Email: jinggzh@sun5.ibp.ac.cn)

Abstract: An efficient fusion expression system with a small fusion partner, His_6-tagged N-terminal fragment of staphylococcal nuclease R, has been constructed and tested with two genes. The results show that the system is not only suitable for overexpression of small proteins and peptides but simplifies purification of target proteins and peptides. The study also provides a practical method for preparation of isotope-labeled protein sample for NMR analysis.

Key Words: protein expression; small fusion partner; NMR sample preparation

INTRODUCTION

Expression of foreign proteins within *E. coli* cell in M9 medium is still a routine method to get enough isotope-labeled samples for NMR analysis. Fusion expression systems, such as fusion with glutathione S-transferase (GST) and maltose-binding protein (MBP), are widely utilized to increase the production and simplify the purification. However, most of these systems have a common disadvantage in that the relatively big fusion partner will reduce the net production of target proteins, especially in the case of small peptides.

We describe here an efficient fusion expression system based on the of T7 promoter, in which a small His_6-tagged N-terminal fragment of staphylococcal nuclease R (HR52), that consists of 58 residues and can be highly and stably expressed in *E. coli* cells [1,2], was chosen as the fusion partner. To examine the expression efficiency of the system, two genes, which encode a novel programmed cell death 5 (PDCD5)

protein of 125 residues [3,4] and a myotoxin α (MYO) peptide of 42 residues [5] were N-terminally fused with the fusion partner. There is a thrombin (LVPR↓GS) or a hydroxylamine (N↓G) cleavage site between the fusion partner and target protein for removing the fusion partner. It is worth noting that PDCD5 protein was directly expressed with very low yield in M9 medium, and myotoxin α was hardly expressed even if it was fused with GST fusion partner in LB medium in our previous experiments (unpublished observation). The results show that this fusion expression system not only allows target proteins to be highly expressed but also simplifies purification of the isotope labeled target proteins for NMR analysis.

MATERIALS AND METHODS

Bacterial strains and plasmids

E. coli DH5α and BL21(DE3) were used as the cloning and expression host cells, respectively. pBV-3 [6], pMT4TFAR19 [4] and pUC-MYO containing the coding sequence for myotoxin α that were used as PCR templates for amplifying fusion partner HR52, PDCD5 and myotoxin α gene respectively. pUC18 was used as the vector for cloning the fusion protein genes and DNA sequencing. pET-3d was kindly provided by Dr. Studier as the vector for expression of fusion proteins.

DNA MANIPULATIONS

Restriction endonucleases and T4 DNA ligase were used as recommended by the manufacturer (Promega). Plasmid isolation and transformation of $CaCl_2$-treated *E. coli* cells were carried out by the procedure described by Sambrook et al. [7]. The fusion partner HR52 was amplified from pBV-3 using two primers: primer1 (18mer, 5'-AT**GAATTCCATGG**GCCAC-3' containing *Eco*RI and *Nco*I sites) and primer 2 (26mer, 5'-TA**GGTACC**CTCTACACGTTTTTTAGG-3' *Kpn*I tail with a TA clamp at 5' end) were used as the forward and reverse primers, respectively. The amplified HR52 sequence was digested with *Eco*RI and *Kpn*I enzymes, and then ligated into *Eco*RI/*Kpn*I digested pUC18 and this plasmid was designated pUC18-HR52. Primer 3 [42mer, 5'-GG**GGTACC**CTGGTTCCGCGTGGTTCCGCGGACGAGGAGCTTG-3', *Kpn*I tail with a GG clamp at the 5' end, followed by a thrombin cleavage site (LVPR↓GS) coding sequence, (in box)] and primer 4 (32mer, 5'-AA**GGATCC**ATTAATAATCGTCATCTTCATCAG-3', *Bam*HI tail with a AA clamp at 5' end) were used as the forward and reverse primers, respectively, to amplify the PDCD5 coding sequence using pMT4TFAR19 as the template. The PCR product was digested with *Kpn*I and *Bam*HI and the PDCD5 coding sequence was ligated with T4 DNA ligase into *Kpn*I and *Bam*HI sites of pUC18-HR52 to yield PDCD5 coding sequence fused to HR52 gene in frame, and this was designated pUC18-HR52-PDCD5. After transformation of the plasmid into *E .coli* DH5α, the clone was validated by DNA sequencing. The target gene encoding myotoxin α was also amplified from pUC-MYO by using the same procedure as described above except that a hydroxylamine cleavage site N↓G was introduced between the fusion partner and myotoxin α instead of the thrombin cleavage site. Finally, the expression plasmid pET-HR52-PDCD5 or pET-HR52-MYO was obtained by insertion of the HR52-PDCD5 or HR52-MYO fragment into the *Nco*I and *Bam*HI sites of pET-3d vector.

Expression, labeling and purification of HR52-PDCD5

A fresh BL21(DE3) clone harboring pET-HR52-PDCD5 was inoculated in 200 ml of LB medium containing 100 μg/ml ampicillin and cultured at 37°C until the optical density (A600) reached 0.8. The whole culture was then transferred to 1 L of the same medium and cultured at 37°C until optimal density (A600) reached 0.6. After centrifugation, the cell pellet was inoculated in 1 L of pre-warmed M9 medium except substituting the $^{15}NH_4Cl$ for NH_4Cl and/or ^{13}C glucose for glucose and cultured at 37°C for 0.5 h, and then IPTG was added to a final concentration of 0.4 mM. After induction at 37°C for 4 h, the cell pellet was collected by centrifugation at 6000 g for 20 min at 4°C, and then resuspended in 50 ml of buffer A (50 mM Tris-HCl, pH 8.0, 500 mM NaCl), followed by ultrasonication on ice for 1 min ×5 to break the cells. The supernatant (45 ml) was recovered after the lysed bacteria were centrifuged at 37,000 g at 4°C for 30 min and directly applied onto a 5×1 cm^2 of chelating Sepharose Fast Flow column at a flow rate 1.0 ml/min, which was charged with Ni^{2+} prepared according to the producer's protocol (Amersham Biosciences) and equilibrated with the same buffer. The column was washed first by a 50 ml of buffer A followed by a 50 ml of buffer B (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 25 mM Imidazole). Finally the column was eluted with buffer C (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 250 mM Imidazole), and monitored by absorbance at 280 nm. The major peak of protein was collected and this was followed by dialysis against distilled water at 4°C. The protein was then lyophilized and stored at -20°C.

The protein samples from each step as described above were analyzed by using 10% SDS-PAGE according to the Blackshear procedure [8], which is a minor modification of the SDS-Tris-glycine system of Laemmli [9]. Proteins were visualized by staining with Coomassie brilliant blue R250.

Thrombin Cleavage of HR52-PDCD5 and Purification of PDCD5

To cleave out the fusion protein HR52, 40 mg of HR52-PDCD5 was dissolved in 20 ml of buffer D (25 mM Tris-HCl, pH 8.0, 2.5 mM $CaCl_2$, 10 mM $MgCl_2$), then thrombin (Sigma T4648) was directly added into the protein solution in the ratio of 10 units per mg of the fusion protein. The digestion was carried out at 4°C for 6 h. The digestion mixture was loaded at a flow rate of 0.2 ml/min onto a 1 ml FPLC Resource Q column (Amersham Biosciences) pre-equilibrated with buffer E (25 mM Tris-HCl, pH 8.0). Chromatography was performed with a linear gradient from 0 to 100% buffer F with buffer E (buffer F, 25 mM Tris-HCl, pH 8.0, 1.0 M NaCl) at room temperature. The major peak (around 0.15 M NaCl gradient) was identified as the target protein PDCD5. The protein was collected and dialysed against distilled water at 4 ^oC. The protein was lyophilized and stored at –20 ^oC.

Expression and purification of myotoxin α

For expression of HR52-myotoxin α, bacterial culture and induction were carried out in the routine way in LB medium. Soluble HR52-myotoxin α fraction was also purified by using Ni^{2+}-chelating affinity chromatography as described above. 38 mg of HR52-myotoxin α was cleaved with hydroxylamine by the method described by Mokes et al [10]. Myotoxin α was purified through second run of Ni^{2+}-chelating affinity chromatography.

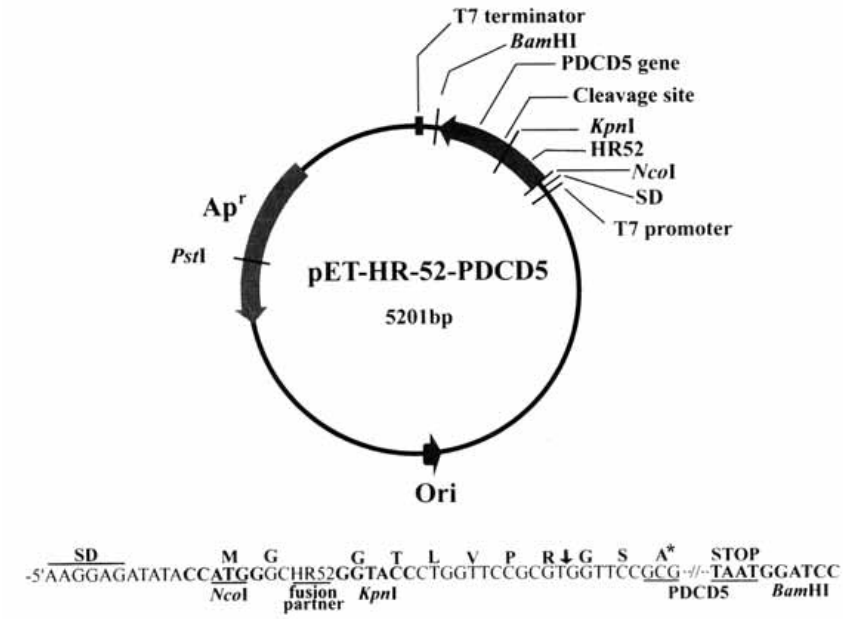

Figure 1. The map of plasmid pET-HR52-PDCD5 for expression of fused HR52-PDCD5 protein. The nucleotide sequence indicates the SD sequence (overlined), the initiator codon and stop codon for the fusion protein underlined (in bold), and shows how is the HR52 fused to PDCD5. The thrombin cleavage site and the first residue of PDCD5 are indicated by ↓ and *, respectively. The construction of pET-HR52-MYO is the same as the pET-HR52-PDCD5 except that the cleavage site is N↓G.

RESULTS AND DISCUSSION

Construction of fusion expression vector for expression of PDCD5 and myotoxin α

As described in Materials and Methods, to construct the fusion expression vector, the fusion partner coding sequence, HR52, was inserted into pUC18 with *Eco*RI and *Kpn*I cut sites to create pUC18-HR52 first. The target gene, PDCD5 or myotoxin α, was then inserted into the pUC18-HR52 with *Kpn*I and *Bam*HI cut sites to yield pUC18-HR52-PDCD5 or pUC18-HR52-MYO. Finally, the DNA fragment encoding HR52-PDCD5 or HR52-MYO cut from the above plasmids with *Nco*I and *Bam*HI enzymes were inserted into pET-3d with the same cut sites to obtain pET-HR52-PDCD5 or pET-HR52-MYO for fusion expression of the target genes. As shown in Figure 1, the fused target gene (HR52-PDCD5 or HR52-MYO) is under the transcriptional control of T7 promoter, and the strong T7 transcription termination is placed downstream of the genes. The HR52-PDCD5 and HR52-MYO genes are placed in the correct reading frame with the initiation codon (ATG) supplied by pET-3d vector. The initiation codon of the fusion protein genes is within 7 bp from the SD sequence. The thrombin or hydroxylamine cleavage site, which will be used to prepare PDCD5 and myotoxin α by using thrombin and hydroxylamine cleavage of the fusion proteins, respectively, is placed between the fusion partner and target protein.

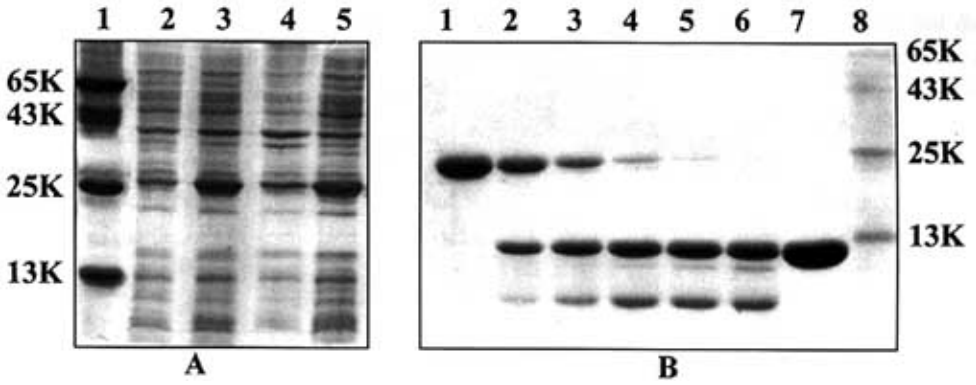

Figure 2. (A) SDS-PAGE analysis of fusion expression of HR52-PDCD5. Lane 1, molecular weight standards; lane 2, whole cell lysate before IPTG induction; lane 3, whole cell lysate after IPTG induction; lane 4, pellet fraction after ultrasonication; lane 5, soluble fraction after ultrasonication. (B) SDS-PAGE analysis of HR52-PDCD5 fusion protein cleaved with thrombin for different times and the purified PDCD5 protein. Lane 1, purified HR52-PDCD5 protein with Ni^{2+}-chelating affinity chromatography; lane 2-6, cleavage of HR52-PDCD5 with thrombin for 1, 2, 4, 6 and 8 h, respectively; lane 7, PDCD5 protein purified by using FPLC Resource Q column; lane 8 molecular weight standards.

Analysis of expression of the HR52-PDCD5 and HR52-myotoxin α in *E. coli* BL21 (DE3)

Expression of the fusion genes in *E. coli* was analyzed by SDS-PAGE. Fig.2A shows the 10% SDS-PAGE pattern of the cell lysate from *E. coli* cells harboring pET-HR52-PDCD5 before and after induction with IPTG. Compared with the uninduced cells (Fig.2A, lane 2), an intensely staining protein band can be observed in the induced total cell lysate (Figure 2A, lane 3), which constitutes more than 25% of the total cellular proteins of the bacterial cell as analyzed by scanning the protein density of the gel. Note that most of HR52-PDCD5 fusion protein is in the supernatant of the cell lysate (Figure 2A, lane 5), indicating that most of the desired recombinant protein is expressed in soluble form. The expression pattern of HR52-myotoxin α was also analyzed by using SDS-PAGE. As shown in Figure 3A, HR52-myotoxin α was highly expressed in *E. coli* cells, and constitutes more than 40% of the total cellular proteins of the bacterial cell (Figure 3A, lane 3). About 40% of the desired fusion protein is expressed in soluble form (Figure 3A, lane 5). It is worth noting that myotoxin α gene could not be expressed when it was fused with GST using pGEX expression vector (unpublished observation). The results indicate that the fusion expression system is suitable for fusion expression of small peptides and proteins.

Simplifying purification of target proteins

Since the fusion partner in the expression vector is a His_6-tagged N-terminal fragment of staphylococcal nuclease R (HR52), fusion proteins can be first purified by using a routine metal-chelating affinity chromatography procedure. As shown in Figure 2B (lane 1) and Figure 3A (lane 6), HR52-PDCD5

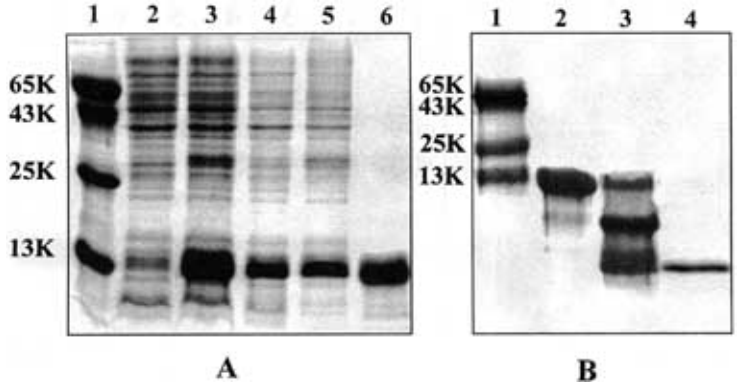

Figure 3. (A) SDS-PAGE analysis of fusion expression of HR52-myotoxin α. Lane 1, molecular weight standards; lane 2, whole cell lysate before IPTG induction; lane 3, whole cell lysate after IPTG induction; lane 4, pellet fraction after ultrasonication; lane 5, soluble fraction after ultrasonication; lane 6, purified HR52-myotoxin α by using Ni^{2+}-chelating affinity chromatography. (B) SDS-PAGE analysis of hydroxylamine cleavage of HR52-myotoxin α in the presence of 3 M GuHCl. Peptide electrophoresis was carried out according to the method described by Schägger et al [11]. Lane 1, molecular weight standards; lane 2, purified HR52-myotoxin α as shown in Figure 3A, lane 6; lane 3, hydroxylamine cleavage of HR52-myotoxin α at 45°C for 4.5 h in the presence of 3 M GuHCl; lane 4, purified myotoxin α by using Ni^{2+}-chelating affinity chromatography.

and HR52-myotoxin α were purified to apparent homogeneity by one-step purification by using Ni^{2+}-chelating affinity chromatography. 40mg of purified HR52-PDCD5 and 38 mg of purified HR52-myotoxin α can be obtained from 1L of M9 and LB bacterial cultures, respectively. Figure 2B shows the cleavage of HR52-PDCD5 with thrombin and purification of PDCD5 on an Amersham FPLC-Resource Q column. The results indicate that thrombin cleavage of HR52-PDCD5 can be completed in 6 h at 4 ^{o}C (Figure 2B lane 5). The main peak that eluted at 0.15M NaCl in the gradient is the purified PDCD5 as shown by SDS-PAGE (Figure 2B, Lane 7). About 20 mg homogeneous PDCD5 can be obtained with a recovery of approximate 70%. Myotoxin α can be purified through a second run of Ni^{2+}-chelating affinity chromatography after hydroxylamine cleavage. 15 mg of purified myotoxin α can be obtained from the soluble fraction of bacterial cell lysate (data not shown). In order to recover the insoluble fraction of HR52-myotoxin α, we tried to use a denaturing lysis buffer (0.1 M phosphate buffer, pH 8.0, 6 M GuHCl) to lyse bacterial cells, and then run a Ni^{2+}-chelating affinity chromatography at the denaturing conditions. The main peak was obtained by using elution buffer (0.1 M phosphate buffer, pH 4.5, 6 M GuHCl). 100 mg of HR52-myotoxin α can be obtained. However, hydroxylamine cleavage in denaturing solution is less efficient and the recovery rate of myotoxin α is low as shown in Figure 3B.

In summary, this study not only provides an efficient fusion expression system for small protein and peptide expression in *E. coli* cells but also provides a practical method for preparation of isotope

labeled protein samples for NMR analysis. Any gene can be inserted into the vector with *Kpn*I and *Bam*HI sites after amplification with appropriate PCR primers. Target proteins or peptides can be obtained after cleavage of fusion proteins with thrombin or hydroxylamine, which are purified by using one-step metal chelating affinity chromatography. In addition, to suppress the basal expression of the fusion protein, the *Nco*I-*Bam*HI cassette (Figure 1) can be also inserted into other pET vector, such as pET-15b and pET-28a [12] using the same restriction enzyme sites. Thus the basal expression of the fusion genes is minimized by the binding of the lac repressor encoded by the lacI gene, to the lac operator immediately downstream of the T7 promoter.

ACKNOWLEDGEMENTS

This work was supported by a grant (Nos. G1999075608 and 30170210) from the China Committee for Science and Technology.

REFERENCES

[1] Jing, G.Z., Liu, L.J., Liu,.Z.G., Zou, Q and Jiang, M.Y. (**1995**) *Progress in Natural Science*, *5*, 569-575.

[2] Zhou, B., Tian, K.G. and Jing, G.Z. (**2000**) *Protein Eng.*, *13*, 35-39.

[3] Liu, H.T., Wang, Y.G., Zhang, Y.M., Song, Q.S., Di, C.H., Chen, G.H., Tang, J. and Ma, D.L. (**1999**) *Biochem. Biophys. Res. Commun.*, *254*, 203-210.

[4] Chen, Y.U., Sun, R.H., Han, W.L., Zhang, Y.M., Song, Q.S., Di, C.H. and Ma, D.L. (**2001**) *FEBS Lett.*, *509*, 191-196.

[5] Nedelkov, D., O'Keefe, M.P., Chapman, T.L. and Bieber, A.L. (**1997**) *Biochem. Biophys. Res. Commun.*, *241*, 525-529.

[6] Li, Y.H., Li, Z.J. and Jing, G.Z. (**1997**) *Biotechnology Technique*, *11*,729-732.

[7] Sambrook, J., Fritsch, E.F. and Maniatis, T. (**1989**) *"Molecular Cloning: A laboratory Manual"* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

[8] Blackshear, P.J. (**1984**) *Methods Enzymol.*, *104*, 237-255.

[9] Laemmli, U. K. (**1970**) *Nature(London)*, *227*, 680-685.

[10] Moks, T., Abrahmsen, L., Holmgren, E., Bilich, M., Olsson, A., Uhlen, M., Pohl, G. Sterky, C. Hultberg, H., Josephson, S., Hollmgren, A., Jornvall H. and Nilsson, B. (**1987**) *Biochemistry*., *26*, 5239-5244.

[11] Schägger, H. and Von Jagow, G. (**1987**) *Anal. Biochem.*, *166*, 368-379.

[12] Studier, F.W. (**1990**) *Methods Enzymol.*, *185*, 60-89.

Received on October 25, 2002, accepted on December 10, 2002.