Title
Screening for the expression of soluble recombinant protein in E. coli

Author
Doyle, Sharon A.

Publication Date
2004

Screening for the Expression of Soluble Recombinant Protein in *E. coli*

Sharon A. Doyle

Running title: Expression Screening for Soluble Proteins

Keywords: protein expression screen, hexahistidine tag, dot blot

Sharon A. Doyle, Ph.D.
DOE Joint Genome Institute
2800 Mitchell Drive
Walnut Creek, CA 94598
Tel: 925-296-5694
Fax: 925-296-5850
Email: sadoyle@lbl.gov

1. Introduction

Protein expression and purification have traditionally been time-consuming, case-specific endeavors, and are considered to be the greatest bottlenecks in most proteomics pipelines (***1***). *Eschericia coli* (*E. coli*) is the most convenient and cost-effective host, although optimal conditions for the expression of different proteins vary widely. Proteins vary in their structural stability, solubility and toxicity in this environment, resulting in differing rates of protein degradation, formation into insoluble inclusion bodies and cell death, thus affecting the amount of soluble protein that can be obtained from *E. coli* grown in culture. A variety of *E. coli* strains designed to address many of these problems have been made, by allowing for disulfide bond formation in the *E. coli* cytoplasm, more efficient usage of codons that are rarely found in *E. coli* genes, and a reduction of specific *E. coli* proteases (***2, 3***). The T7 expression system (***4-6***) has also helped to alleviate many problems associated with recombinant expression by allowing for tightly regulated induction of expression. However variability in the recombinant expression of proteins still exists. Induction strength and time, and growth temperature are just a few parameters that can have drastic effects on the amount of soluble protein that can be obtained from the *E. coli* host. Adding further complexity to this problem, affinity tags that are used as a means to standardize protein purification have varying effects on protein stability and solubility (***7***). The hexahistidine tag, which allows for protein to be adsorbed to Ni^{2+}-charged resin (***8***), is often used because it is small and often does not need to be removed by protease digestion prior to downstream experimentation. Another

useful tag is the maltose binding protein (MBP) tag, which allows for the protein to be captured with amylose resin (***9***), and can stabilize and solubilize the fusion partner (***10***). A variety of affinity tags, in addition to their placement at the N- or C- terminal end of the coding sequence, needs to be tested for optimal protein expression.

To take full advantage of the variety of strategies developed to improve the expression of soluble protein in *E. coli*, a means to test easily and rapidly many growth parameters is necessary. This chapter describes a dot-blot expression screen to test the effects of growth and induction parameters on the yield of soluble protein. The expression screen is used to detect hexahistidine-tagged proteins expressed in *E. coli*, however it is adaptable for the detection of other affinity tags or fusion partners that have suitable antibodies available. In this example, induction time and temperature are tested, however, it can be used to test additional parameters such as affinity tag type and placement, *E. coli* host type, and growth medium formulations. Results of the screen may be used to segregate samples into groups for further parallel processing in an efficient high-throughput protein production pipeline.

2. Materials

1. Rosetta (DE3) (Novagen, Madison, WI) and BL21 Gold (DE3) (Stratagene, La Jolla, CA) *E. coli* competent cells or other expression host strain
2. cDNA clones in T7 expression vector
3. LB media and agar plates

4. Chloramphenicol

5. Kanamycin

6. 24-well (7 ml) round bottom blocks

7. IPTG (isopropyl-beta-D-thiogalactopyranoside)

8. Lysis buffer: 50 mM sodium phosphate buffer (pH 8.0), 300 mM NaCl, 10 mM imidazole. Adjust pH to 8.0 and store at 4 °C.

9. Lysozyme

10. Benzonse (Novagen, Madison, WI) or DNAse

11. Protease inhibitor cocktail (for histidine-tagged proteins) (Sigma-Aldrich, St. Louis, MO)

12. Phenylmethylsulfonylfluoride (PMSF)

13. Protran nitrocellulose membrane (0.2 micron pore size) (Schleicher and Schuell, Keene, NH)

14. TBS: 6 mM Tris, 150 mM NaCl, pH 7.5

15. TBS T/T: 20 mM Tris, 500 mM NaCl, 0.05% Tween 20, 0.2% Triton X-100, pH 7.5

16. Blocking buffer: TBS with 3% bovine serum albumin (BSA)

17. Penta His HRP conjugate kit (Qiagen, Valencia. CA)

18. Metal enhanced DAB substrate kit (Pierce, Rockford, IL)

3. Methods

3.1 Protein Expression and Cell Lysis

The following protocol describes the expression of proteins using the pNHis vector (***11***) in the *E. coli* strains Rosetta (DE3) pLysS and BL21 Gold (DE3), based on previous results (not shown). The strain Rosetta (DE3), used for the proteins of *Ciona intestinalis*, is optimized for the expression of proteins containing rare codons often found in genomes enriched with an abundance of either GC or AT base pairs. The strain BL21 Gold (DE3) is used for the expression of proteins from *Xylella fastidiosa*. Note that alternative expression plasmids and *E. coli* strains may be used (*see* **Note 1**) requiring adjustments to the general protocol such as the use of antibiotics.

1. Transform or electroporate Rosetta (DE3) pLysS or BL21 Gold (DE3) cells with pNHis expression plasmids following the manufacturer's protocol. Plate onto LB agar plates containing 50 μg/ml kanamycin and 50 μg/ml chloramphenicol (Rosetta cells only) and grow overnight at 37 °C.

2. Pick a single colony of each sample and grow 5 ml cultures (in a 24-well block) of LB containing 50 μg/ml kanamycin and 50 μg/ml chloramphenicol (Rosetta cells only) under non-inducing conditions (no IPTG) by shaking at 180 rpm at 37 °C overnight.

3. Identify the parameters that will be tested, such as inducer concentration, temperature and time of culture growth following induction of protein expression (*see* **Note 2**). Keep mind that some sets of conditions will require a separate 24-well growth block. Aliquot 5 ml of LB containing 50 μg/ml kanamycin and 50 μg/ml chloramphenicol (Rosetta cells

only) into each well of the blocks, and inoculate with 100 µl of the overnight starter cultures.

4. Grow cultures by shaking at 180 rpm at 37 °C until an optical density (O.D.) at 600 nm wavelength of approximately 0.6 –0.8 is obtained. Add IPTG at the desired concentration and continue to shake at the desired temperature for the desired length of time. In this example growth temperature (18, 25, 30 and 37 °C) and time (4 h or overnight) following induction with 1 mM IPTG were tested.

5. Pellet the cells by centrifugation at 4 °C and freeze at –70 °C or in a dry ice/ethanol bath. Prepare fresh lysis buffer by adding 1 mg/ml lysozyme, 2.5 U/ml Benzonase nuclease, 2 mM $MgCl_2$, 2 µl/ml Protease inhibitor cocktail, 1 mM PMSF. Resuspend the frozen pellets in 500 µl fresh lysis buffer and shake on a plate shaker at 4 °C for 30 min to lyse the cells.

6. Remove 10 µl of crude lysate (spotted on the membrane as total protein) and remove the insoluble material by centrifugation. The supernantant (cleared lysate) and crude lysate are then used in the dot blot procedure described below to determine the amount of soluble and total recombinant protein, respectively.

3.2 Dot Blot Procedure

1. Spot 2 μl of crude and cleared lysate onto Protran nitrocellulose membrane using a multichannel pipettor. Also spot a set of standards. This example spotted a sample protein, isocitrate dehydrogenase (IDH) in a range of 15-1500 ng. Allow the membrane to dry completely.

2. Incubate the membrane by gentle shaking at room temperature as follows: TBS, 5 min x 3, Blocking buffer, 30 min, TBS T/T 5 min x 3, Penta His HRP conjugate (1:1000 in Blocking buffer) 30 min, TBS T/T, 5 min x 5.

3. Dilute the DAB concentrate (10X stock) with stable peroxide buffer to 1X following the manufacturer's protocol. Submerse the membrane in the DAB solution and develop for approximately 3 minutes (**Fig. 1**). Rinse membrane with H_2O, incubate in a dark place until it is completely dry, and scan the image using a flatbed scanner.

4. Notes

1. A wide variety of *E coli* strains designed for protein expression are commercially available, derived from the E. coli strain BL21, that are tailored to facilitate disulfide bond formation, fine tune protein expression levels, enhance the expression of proteins that contain rare codons, and other specific requirements. These strains are available as lambda DE3 lysogens, which carry a chromosomal copy of the T7 RNA polymerase gene under control of the lac promoter (inducible by IPTG). Once induced, the T7 RNA polymerase drives expression of genes that are under control of the T7 promoter, such as those cloned into the pET family of expression vectors (Novagen, Madison, WI).

Dramatic differences in protein expression can be seen when different strains are used for the production of proteins in many cases, therefore it is useful to test several strains when producing many proteins in a high-throughput format without knowledge of their specific characteristics.

2. Conditions most commonly tested are inducer (IPTG) concentration (0.1 mM to 1.0 mM), growth temperature after induction (18 °C to 37 °C), and duration of time following induction (4 h to overnight). Additional parameters to consider testing include *E. coli* strain, media formulation, and placement of affinity tag (N- or C- terminal).

Acknowledgements

The author thanks Jennifer Massi and Shirin Fuller for technical assistance, and Michael Murphy, Peter Beernink, and Paul Richardson for reading of the manuscript. This work was performed under the auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, by the University of California, Lawrence Livermore National Laboratory (No. W-7405-Eng-48), Lawrence Berkeley National Laboratory (No. DE-AC03-76SF00098), and Los Alamos National Laboratory (No. W-7405-ENG-36).

References

1. Pedelacq, J.D., Piltch, E., Liong, E.C., Berendzen, J., Kim, C.Y., Rho, B.S., Park, M.S., Terwilliger, T.C., & Waldo, G.S. (2002) Engineering soluble proteins for structural genomics. Nat. Biotechnol. **20**, 927-932.

2. Novagen (Madison, WI)

3. Stratagene, (La Jolla, CA)

4. Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**, 113-130.

5. Rosenberg, A.H., Lade, B.N., Chui, D., Lin, S., Dunn, J.J., and Studier, F.W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56**, 125-135.

6. Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Meth. Enzymol. **185**, 60-89.

7. Braun, P. and LaBaer, J. (2003) High-throughput protein production for functional proteomics. Trends in Biotechnology **21** (9), 383-388.

8. Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. & Nordheim, A. (1991) Rapid and efficient production of native histidine-tagged protein expressed by recombinant vaccinia virus. Proc. Natl. Acad. Sci. USA **88**, 8972-8976.

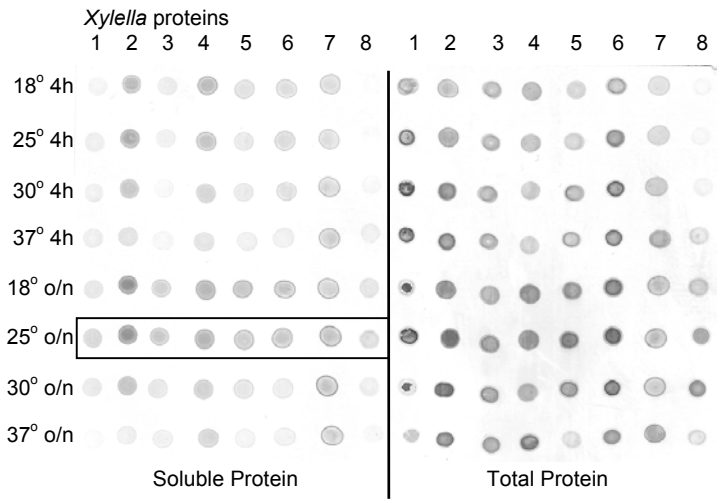
9. Riggs, P. (1992) Expression and purification of maltose-binding protein fusions. In Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) Greene Publishing and Wiley-Interscience, New York. pp. 16.6.1-16.6.14.

10. Kapust, R.B. & Waugh, D.S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci. **8**, 1668-1674.

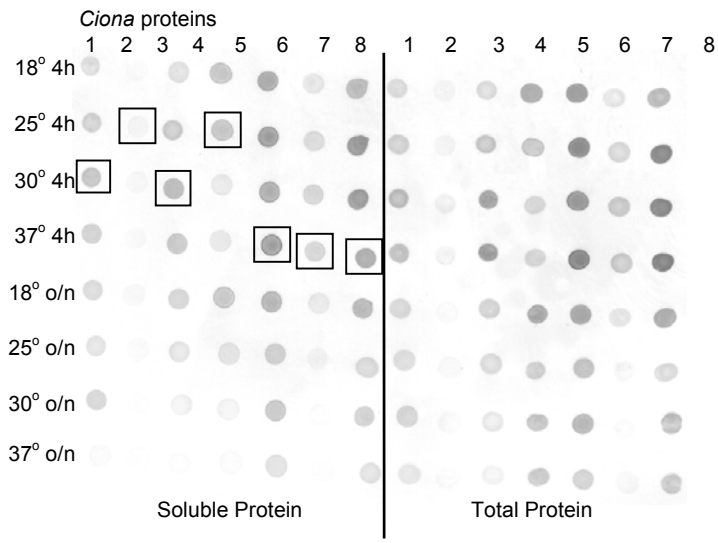
11. Doyle SA, Murphy MB, Massi JM, and Richardson PM (2002) High-Throughput Proteomics: A Flexible and Efficient Pipeline for Protein Production. *J. Proteome Research* Dec. **1** (6): 531-536.

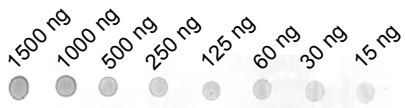
Fig. 1. Expression screen dot blots of A) *Xylella* samples and B) *Ciona* samples, grown for 4 hours or overnight (o/n) at 18, 25, 3, and 37 °C. The standard curve is shown below. The spots outlined with squares indicate samples that would be chosen for protein purification.

A.


Xylella proteins
1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8
18° 4h
25° 4h
30° 4h
37° 4h
18° o/n
25° o/n
30° o/n
37° o/n
Soluble Protein
Total Protein

B.


Ciona proteins
1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8
18° 4h
25° 4h
30° 4h
37° 4h
18° o/n
25° o/n
30° o/n
37° o/n
Soluble Protein
Total Protein


1500 ng
1000 ng
500 ng
250 ng
125 ng
60 ng
30 ng
15 ng