2

Protein Expression Arrays for Proteomics

Michele Gilbert, Todd C. Edwards, and Joanna S. Albala

Summary

As biology approaches the 50th year of deciphering the DNA code, the next frontier toward understanding cell function has protein biochemistry in the form of structural and functional proteomics. To accomplish the needs of proteomics, novel strategies must be devised to examine the gene products or proteins, emerged as *en masse*. The authors have developed a high-throughput system for the expression and purification of eukaryotic proteins to provide the resources for structural studies and protein functional analysis. The long-term objective is to overexpress and purify thousands of proteins encoded by the human genome. This library of proteins—the human proteome—can be arrayed in addressable format in quantities and purities suitable for high-throughput studies. Critical technology involved in efficiently moving from genome to proteome includes parallel sample handling, robust expression, and rapid purification procedures. Automation of these processes is essential for the production of thousands of recombinant proteins and the reduction of human error.

Key Words:

Protein array; baculovirus; insect cell; protein expression; purification; automation; robotics.

1. Introduction

1.1. Overview: Array-Based Proteomics

The key advantage to array-based methods for protein study is the parallel analysis of samples in a high-throughput fashion. Similar to the DNA microarray, this approach requires miniaturization technologies, high sample throughput, and automation. Array-based methods for protein analysis afford a high-throughput format by which to screen protein–protein, protein–DNA, and protein–small molecule interactions and provides important functional information for newly identified genes that are derived from genome projects. Protein arrays hold the potential to identify these interactions as well as provide a means for differential expression and protein profiling between different cell types.

1.2. Generation of Protein Arrays

Proteins, peptides, and antibodies have been analyzed using a microarray format, and protein arrays have been produced using various media and a diversity of immobilization chemistries on surfaces such as nitrocellulose, polyvinylidene fluoride, silicone, glass, and plastic (for review, *see* **refs. 1–5**). Use of a standard glass microscope slide to bind proteins or antibodies provides a cheap, easily manipulated format that is amenable to many chemical modifications, as surface chemistry is critical when preparing protein arrays. Proteins, peptides, or antibodies can be applied to the array surface by ink-jet or contact printing in a similar manner to those used in spotting a DNA array *(6)*. Generally, most analyses use fluorescent or radiolabeled targets for capture by proteins bound to the array, enzymatic or colorimetric analysis for functional assay, and mass spectrometry or surface plasmon resonance for detection.

1.3. Protein Production for Generation of Protein Arrays

The earliest bottleneck to the generation of protein arrays is obtaining large numbers of soluble, purified, functional proteins for direct application onto the array or for the generation of antibodies. Recombinant expression in *Escherichia coli* has become the standard because of robust production, low cost, and ease of use. Several laboratories to date have successfully produced and purified large numbers of proteins using high-throughput strategies in *E. coli* either by recombinant or in vitro means *(7–10)*. To overcome many of the limitations arising from prokaryotic expression, such as formation of inclusion bodies, misfolding of proteins, and lack of posttranslational modifications, several eukaryotic systems have been developed using either yeast, insect, or mammalian cells for host expression. Dual-use methods for recombinant expression of prokaryotic and eukaryotic systems have also been devised as well as cell-free systems to expand recombinant protein production capabilities *(11)*.

Automation is key to providing the throughput needed for proteomic studies involving hundreds to thousands of proteins. Many protein production methodologies lend themselves to robotic manipulation because of the repetitive nature of the procedures, such as plasmid isolation, polymerase chain reaction (PCR), DNA quantitation, cell culture, and affinity purification. The authors have developed an automatable system for high-throughput protein production in baculovirus *(12,13)*. Using complementary DNA (cDNA) clones from the LLNL-I.M.A.G.E. collection *(14)*, they can produce recombinant protein in a miniaturized, high-throughput format to derive large numbers of recombinant proteins for downstream functional applications, such as protein microarrays, antibody production, or pathway reconstitution (**ref. 15**; *see* **Note 1**).

2. Materials

2.1. PCR Production of cDNA Clone Inserts

1. *E. coli* from LLNL-IMAGE cDNA Collection.
2. 96-well round-bottom plates.
3. Luria Bertani (LB) broth/ampicillin/glycerol medium.
4. Cloned *Pfu* polymerase (Stratagene).
5. *Asc*I and *Fse*I enzymes (New England Biolabs).

6. 10X PCR buffer.
7. Deoxynucleotide-triphosphates (dNTPs).
8. QIAquick 96-well PCR purification kit (Qiagen).

2.2. Transfer Vector Design and Ligation of cDNA Inserts

1. pBacPAK9 (Clontech).
2. Shrimp alkaline phosphatase (SAP) (Fermentas).
3. One Shot TOP10 chemically competent *E. coli* (Invitrogen).
4. LB/ampicillin/agar 100-mm plates.
5. Wizard miniprep kit (Promega).
6. LB/ampicillin.

2.3. Transfection and Viral Amplification

1. *Sf9* insect cells.
2. Superfect transfectant (Qiagen).
3. IPL-41 insect cell media.
4. Linearized baculoviral DNA (Baculogold, Pharmingen).
5. SF900II insect cell media (Invitrogen).
6. Fetal bovine serum (FBS).

2.4. Deep-Well Viral Amplification and Protein Expression

1. 96-deep-well plate (Marsh Bioproducts).
2. 2.38-mm stainless steel beads (V& P Scientific).
3. 1% Pluronic F68.
4. Gas-permeable seal (Marsh Bioproducts).
5. Carousel Levitation Magnetic Stirrer (V& P Scienctific).
6. Sorvall RT-6000 centrifuge.

2.5. Protein Purification and Analysis

1. Lysis buffer: 20 mM Tris-HCl pH 8.0, 1 mM ethylene glycol *bis* (2-aminoethyl ether)-N,N,$N'N'$-tetraacetic acid (EGTA), 1 mM MgCl$_2$, 0.5% v/v N-octoglucoside.
2. Microplate mixer MT-360 (TOMY).
3. Sodium chloride.
4. Immunoaffinity beads.
5. Wash buffer: 20 mM Tris-HCl pH 8.0, 1 mM EGTA, 1 mM MgCl$_2$, 100 mM NaCl.
6. Elution buffer: wash buffer plus 5 µg/mL peptide.
7. 96-well filter plate (Whatman, 0.45 µM cellulose acetate filter).
8. Vacuum manifold (Whatman).
9. ECL Plus kit (Amersham).
10. 10% Tris-HCl denaturing gels (Novex).
11. Coomassie blue dye.

3. Methods

The methods developed for miniaturized protein production in baculovirus are described in the following sections. The steps are (a) PCR production of cDNA clone inserts, (b) transfer vector design and ligation of cDNA inserts, (c) transfection and viral amplification, (d) deep-well viral amplification and protein expression, and (e) protein purification and analysis (*see* **Note 2**).
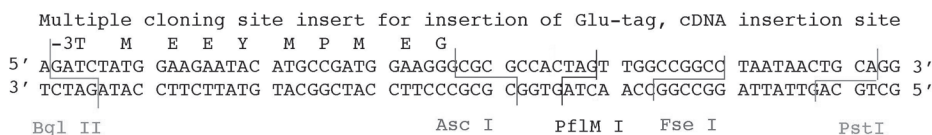
```
  Multiple cloning site insert for insertion of Glu-tag, cDNA insertion site
    -3T   M    E   E  Y    M    P   M    E   G
5' AGATCTATG GAAGAATAC ATGCCGATG GAAGGCGC GCCACTAGT TGGCCGGCC TAATAACTG CAGG 3'
3' TCTAGATAC CTTCTTATG TACGGCTAC CTTCCCGCG CGGTGATCA ACCGGCCGG ATTATTGAC GTCG 5'

  Bgl II                                    Asc I     PflM I   Fse I          PstI
```

Fig. 1. The multiple-cloning site of pMGGlu, which contains the *Asc*I and *Fse*I cloning sites and the MEEYMPMEG (Glu) epitope tag.

3.1. PCR Production of cDNA Clone Inserts

The upstream molecular biology of the baculovirus-based system relies on many of the same techniques that have been applied for production of recombinant proteins in *E. coli.* These methods can also be used to subclone the genes of interest into an appropriate transfer vector for recombination with the baculovirus genome. The authors' scheme for amplification of cDNA clones begins by the generation of 5' gene-specific primers that are paired with a 3' vector-specific primer. The 5' gene-specific PCR primer is designed to contain the rare cutter *Asc*I site, and the 3' vector-specific primer contains a rare cutter *Fse*I site (*see* **Fig. 1**).

1. Aliquot 5 μL *E. coli* containing the cloned genes of interest into 96-well round-bottom plates containing 95 μL LB/ampicillin/glycerol medium and grow overnight at 37°C.
2. Perform PCR directly on a 1:100 dilution of the bacterial cultures using *Pfu* polymerase. The PCR conditions are 96°C for 3 min, 35 cycles of 96°C for 30s, 50°C for 30 s, then 6 min at 72°C, and they have been tested on genes ranging in size from 386 bp to 2409 bp. This cycle is followed by a final extension at 72°C for 10 min.
3. The PCR reaction includes the following: 10X PCR buffer diluted to a final concentration of 1X, dNTPs (25 m*M* each), 0.5 μ*M* final concentration of 5' primer and 3' primer, a 1:100 final dilution of *E. coli* in ddH$_2$O, and 5 U cloned *Pfu* polymerase in a final reaction volume of 50 μL.
4. Purify the PCR products using a Qiagen 96-well format (QIAquick 96 PCR purification kit) and elute into 50 μL ddH$_2$O.
5. Enzymatically digest the resulting PCR products with *Asc*I and *Fse*I, and purify the digested samples with the QIAquick 96 PCR purification kit.

3.2. Transfer Vector Design and Ligation of cDNA Inserts

For the creation of recombinant baculoviruses, a modified transfer vector was designed based on the pBacPAK9 transfer vector from Clontech (*see* **Fig. 2**). A "Glu" immunoaffinity tag *(16)* followed by exonuclease sites for the rare cutters *Asc*I, *Pfl*M*I, and *Fse*I were added between the *Bgl*II and *Pst*I site of the multiple-cloning site of the pBacPAK9 transfer vector to generate the modified transfer vector called pMGGlu.

1. Linearize the pMGGlu vector with *Asc*I and *Fse*I.
2. Dephosphorylate the vector with SAP in preparation for inserting the clones of interest.
3. Ligate each of the clones into the cut and dephosphorylated pMGGlu vector in 96-well format.
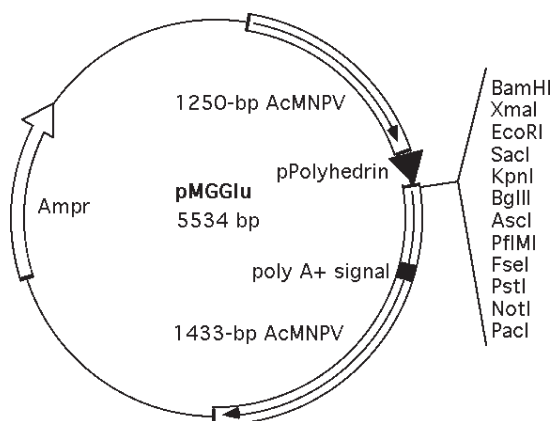
Fig. 2. Schematic diagram of pMGGlu, which is derived from pBAKPAK9 from Clonetech (which contains the Glu Immunoaffinity site followed by the rare cutter sites AscI and FseI for cloning).

4. Inactivate the reaction by heating at 65°C for 10 min.
5. Transform the ligation reactions into TOP10 cells from a One Shot kit, and then plate each transformation onto LB/ampicillin/agar (*see* **Note 3**).
6. Isolate two *E. coli* colonies for each cDNA clone and grow overnight in 3 mL LB/ampicillin.
7. Isolate plasmid DNA using the Wizard miniprep kit (*see* **Note 4**).
8. Screen the plasmid DNA by enzymatic digestion with *Asc*I and *Fse*I followed by agarose gel electrophoresis to determine if the correct size insert for the PCR gene product of interest is contained within the pMGGlu transfer vector.

3.3. Transfection and Viral Amplification

Once the genes of interest are inserted into the baculoviral transfer vector, pMGGlu, the vectors containing the cloned cDNAs are transfected into *Sf9* insect cells along with linearized baculoviral DNA. The cDNA is transferred from the transfer vector to the baculoviral genome by homologous recombination using the cellular machinery of the host insect cell.

1. Place *Sf9* insect cells into a 96-well flat-bottomed tissue-culture plate at 0.5×10^5 cells/ well, and allow the cells to adhere for at least 30 min in a humidified 27°C-incubator.
2. Prepare a 1:50 dilution of SuperFect transfectant in IPL-41 media, and allow the solution to interact for a minimum of 10 min for micelle formation to facilitate transfection.
3. After 10 min, combine 5–10 ng of recombinant transfer vector and 5–10 ng of linearized baculoviral DNA per well, and incubate with the SuperFect solution at a final dilution of 1:100 in IPL-41 media (34 μL transfection cocktail per well) for at least 10 min.
4. Aspirate the media off the cells, and add the transfection cocktail (linearized baculoviral DNA, recombinant transfer vector, and Superfect) to the adherent cells.

5. Allow the cells to transfect for 2–3 h in a humidified 27°C chamber, and then add 70 μL of SF900II media containing 10% FBS to each well.

6. Incubate the cells for 4 d in a humidified 27°C chamber for viral cultivation.

7. After 4 d, plate fresh *Sf9* insect cells onto a new 96-well tissue-culture plate at a density of 2×10^4 cells/well in 70 μL of SF900II media, and allow the cells to adhere for 30 min.

8. After the cells adhere, add 30 μL of supernatant (containing the recombinant baculoviral particles that had been successfully created from the original transfection plate) to each well of newly plated cells.

9. Continue viral amplification for 4 d.

10. Repeat amplification **steps 7–9** in 96-well format two to four more times.

3.4. Deep-Well Viral Amplification and Protein Expression

The final round of viral amplification is performed in a 96-deep-well plate (2 mL) to generate a larger volume of virus for protein production. A Carousel Levitation Magnetic Stirrer is used to culture up to 12 96-deep-well plates at once, for a total of 1152 clones to be produced simultaneously.

1. Add a 2.38-mm steel ball to each well in the 96-deep-well plate, and then add 1.5 mL of *Sf9* insect cells at a density of 1.5×10^6 cells/mL in SF900II media containing 1% Pluronic F68 to each well.

2. Add virus at 5–10% v/v to the cells and cover the 96-deep-well plate with a gas-permeable seal.

3. Incubate the cells for 4 d on a carousel stirrer at a speed setting of 50 at 27°C.

4. Harvest the cells by centrifugation at 3000*g* on a Sorvall RT-6000.

5. Retain the supernatant containing the recombinant virus and discard the cell pellet.

6. For protein production, repeat **steps 1–3**, but only incubate the cells for 48 h rather than 4 d.

7. Harvest the cells by centrifugation at 3000*g* on a Sorvall RT-6000.

8. Aspirate the supernatant and freeze the cell pellet overnight at –80°C.

3.5. Protein Purification and Analysis

Protein purification from insect cells proceeds in a similar fashion to that of other cell types. Various affinity chromatographic techniques are available for protein purification. This method employs immunoaffinity chromatography by use of an antibody conjugated to a Sepharose matrix. The antibody was generated against the Glu peptide epitope tag *(16)*.

1. Thaw the frozen cell pellets and add 0.5 mL lysis buffer to each well of the 96-deep-well plate, leaving the stainless steel balls in the wells to aid in mechanical lysis.

2. Shake the plate on a Microplate mixer MT-360 (TOMY) for 10 min at room temperature to resuspend and lyse the cells.

3. Add NaCl to each sample to a final concentration of 100 m*M*, and shake the plate for an additional 5 min.

4. Centrifuge the lysate at 3000*g* for 20 min.

5. Place 100 μL of the immunoaffinity column matrix in a 96-deep-well plate.

6. To equilibrate the matrix, wash two times by adding 500 μL wash buffer, gently agitate, and centrifuge at 1000*g* for 10 min.

7. Transfer the supernatants containing the soluble protein onto the immunoaffinity matrix, and save the insoluble cell pellets for future examination.
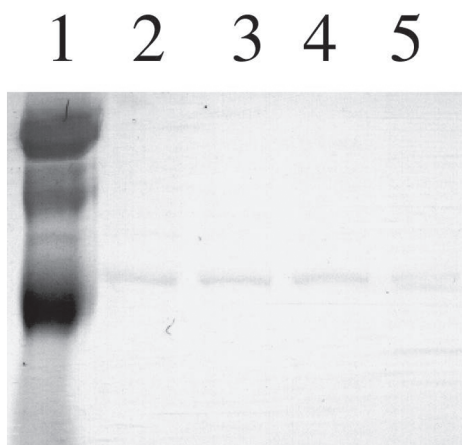
Fig. 3. Coomassie blue staining of purified Gus separated by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). *Lane 1*: Kaleidoscope markers (Bio-Rad). *Lanes 2–5*: 30 μL of purified Gus from four individual clones. (*See* **Note 2**.)

8. Bind proteins to the matrix for 10 min with gentle agitation by pipet.
9. Centrifuge the matrix at 1000*g* for 10 min, and carefully remove the supernatant.
10. Wash the matrix two times by adding 500 μL wash buffer, gently agitate, and centrifuge at 1000*g* for 10 min.
11. After discarding the supernatant, centrifuge at 1200 *g* for 5 min.
12. Discard any remaining supernatant.
13. Resuspend the matrix in 100 μL elution buffer and transfer to a 96-well filter plate (Whatman, 0.45 μ*M* cellulose acetate filter).
14. Allow the elution buffer to interact with the beads for 5 min.
15. Apply light vacuum to collect the supernatant containing the eluted protein in a fresh 96-well collection plate.
16. Analyze the soluble and insoluble protein fractions by gel electrophoresis and Western blot analysis.
17. Detect protein with an enhanced chemiluminescence (ECL) Plus kit.
18. Estimate protein purity by gel electrophoresis followed by Coomassie blue staining.

4. Notes

1. Because the procedures are performed in a 96-well format, many of the processes described can be automated using standard liquid-handling robots. A robust database is critical to track each cDNA clone through the many processes to produce a purified protein. Future iterations of the protocols will be implemented as modules for (a) PCR production of cDNA clone inserts; (b) ligation of cDNA inserts; (c) transfection and viral amplification; (d) viral amplification and protein expression; and (e) protein purification and analysis on these robots with Web-based graphic interface to access the database.
2. Throughout production, the gene for β-glucoronidase was used as a control. The efficacy of transfection, infection, and protein production can be measured by examining the abil-

ity of this enzyme to break down its substrate X-Glucuronide, which results in a blue-colored product that can be quantified by spectrophotometric analysis at 630 nm. An example of this purified protein is shown in **Fig. 3**.

3. Originally, the authors anticipated that the ligation reaction could be directly transfected into the insect cells along with linearized baculoviral DNA, to avoid the *E. coli* transformation step. However, it was determined that the low probability of ligation (approx 100 clones per transformation) resulted in decreased transfection efficiency. Therefore, the subcloning into *E. coli* was necessary to increase the probability of a productive homologous recombination event.

4. Although the transformation and DNA isolation were performed offline and not in 96-well format, kits and plates do exist to perform these steps in an automatable, 96-well format (Promega, Qiagen).

Acknowledgments

References

1. Holt, L. J., Enever C., deWildt R. M. T., et al. (2000) The use of recombinant antibodies in proteomics. *Curr. Opin. Biotechnol.* **11,** 445–449.

2. Zhu, H. and Synder, M. (2001) Protein arrays and microarrays. *Curr. Opin. Chem. Biol.* **5,** 40–45.

3. Cahill, D. J. (2001) Protein and antibody arrays and their medical applications. *J. Immunol. Methods* **250,** 81–91.

4. Reineke, U., Volkmer-Engert, R., and Schneider-Mergener, J. (2001) Applications of peptide arrays prepared by the SPOT-technology. *Curr. Opin. Biotechnol.* **12,** 59–64.

5. Albala, J. S. (2001) Array-based proteomics: the latest chip challenge. *Expert Rev. Mol. Diagn.* **1,** 145–152.

6. Schena, M., Shalon, D., Heller, R., et al. (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. USA* **93,** 10,614–10,619.

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

8. Larsson, M., Graslund, S., Yuan, L., et al. (2002) High-throughput protein expression of cDNA products as a tool in functional genomics. *J. Biotechnol.* **80,** 143–157.

9. Bussow, K., Cahill, D., Nietfeld, W., et al. (1998) A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library. *Nucleic Acids Res.* **26,** 5007–5008.

10. Braun, P., Hu, Y., Shen, B., et al. (2002) Proteome-scale purification of human proteins from bacteria. *Proc. Natl. Acad. Sci. USA* **99,** 2654–2659.

11. Lueking, A., Holz, C., Gotthold, C., et al. (2000) A system for dual protein expression in *Pichia pastoris* and *Escherichia coli*. *Protein Expr. Purif.* **20,** 372–378.

12. Gilbert, M. and Albala, J. S. (2002). Accelerating code to function: sizing up the protein production-line. *Curr. Opin. Chem. Biol.* **6,** 102–105.

13. Albala, J. S., Franke, K., McConnell, I. R., et al. (2000) From genes to proteins: high-throughput expression and purification of the human proteome. *J. Cell. Biochem.* **80,** 187–191.
14. Lennon, G., Auffray, C., Polymeropoulos, M., et al. (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33,** 151–152.
15. MacBeath, G. and Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput functional determination. *Science* **289,** 1760–1763.
16. Rubinfeld, B., Munemitsu, S., Clark, R., et al. (1991) Molecular cloning of a GTPase activating protein specific for the Krev-1 protein p21rap1. *Cell* **65,** 1033–104.

Springer