Mammalian Expression Cloning of Nucleic Acid Binding Proteins by Agarose Thin-Layer Gelshift Clone Selection

BioTechniques 33:868-872 (October 2002)

A. Dobi, W. Debnam, C. Dalgard, A. Owusu, and D.v. Agoston

Uniformed Services University,
Bethesda, MD, USA

ABSTRACT

Here we report a functional screening technique to identify cDNAs encoding mammalian nucleic acid binding proteins. We have combined cDNA expression cloning with the agarose thin-layer gelshift assay technique to detect specific nucleic acid binding proteins from a mammalian expression library. We divided this cDNA expression library into multiple pools and transfected mammalian cells with the individual pools. Following transfection, we tested the expressed proteins for DNA-binding activity by agarose thin-layer electrophoretic gelshift assay. After we identified a single expression pool for the presence of a DNA-binding protein, the corresponding cDNA pool was further divided into smaller aliquots. Then, the cDNA expression and gelshift clone selection was repeated until a single clone was isolated. In contrast to traditional polyacrylamide gels, the agarose thin-layer is significantly faster and resolves larger DNA-protein complexes. This method can be widely used for the cDNA cloning of DNA- and RNA-binding proteins from various mammalian host cells.

INTRODUCTION

Expression cloning has been used for functional gene identification in laboratories with diverse research interests. For expression cloning, cells are transformed with vectors containing cDNA sequences coding for various proteins. Expressed proteins then are assayed for specific biological functions, enabling the selection of a specific cDNA clone. RNA- or DNA-binding proteins can also be identified by expression cloning. The one-hybrid system is a widely used clone selection method for DNA- or RNA-binding activity (6). This cloning technology employs indirect detection of protein-DNA interactions by the trans-activation of a reporter gene by fusion hybrid proteins. The one-hybrid systems, however, do not permit a free choice of host cell for protein expression. Consequently, post-translational processing by one-hybrid host cells may not be optimal for detecting the activity of a DNA-binding protein. In contrast to the conserved transcription and translation process, post-translational modifications can significantly vary between species and cell types. A clone selection method called sib (subdivision of positive pools) technology (1) offers a free choice of host cells for cDNA expression. Sib technology is a method of sequential division of a heterogeneous, high-complexity gene library into smaller, low-complexity pools until a single clone is identified by a biological assay. However, this assay system is limited by the sensitivity of the employed biological assay (7). To exceed the detection limit, it is necessary to di-

vide the expression library into numerous pools to lower the clone complexity in the divided pools. If too many pool divisions are necessary, then the number of assays to perform may exceed the laboratory scale. For DNA-binding proteins, a commonly used test is the polyacrylamide gel-based electrophoretic mobility shift assay (4). Abundant transcription factors have been cloned by the combination of sib and polyacrylamide gelshift techniques (8). However, the application of polyacrylamide gels for sib selection failed to become a popular approach. However, the in vitro expression cloning technique (IVEC) has emerged as a powerful application of the sib principle (5). This method is based on the effective in vitro translation of proteins. However, natural post-translation modifications are not optimal in IVEC systems.

We have reasoned that the sib principle is advantageous for in vivo mammalian expression cloning of DNA-binding proteins because a favorable mammalian host cell can be chosen for the expression of a mammalian cDNA library. This expression host would provide the proteins with proper phosphorylation, glycosylation, ubiquitination, cleavage, folding, co-factors, chaperons, and interaction partners to gain complete DNA-binding activity. However, clone selection by polyacrylamide gel electrophoresis did not appear feasible for two reasons. First, the small pore size of polyacrylamide gels does not allow larger protein-DNA complexes to enter the gel matrix during electrophoresis. Second, the lengthy processing time of polyacrylamide gelshift assays makes it inefficient. To get

around these obstacles, we have combined the *sib* cDNA expression technique with agarose thin-layer electrophoresis for DNA-binding assay. Large protein-DNA complexes, characteristic of mammalian transcription factors, can be separated on agarose gels in a fast and efficient manner. The faster screening procedure allowed us to divide the cDNA library into more pools. More pools resulted in lower clone complexity in the individual pools (7). By lowering the clone complexity in the pools, detection of low-abundance transcription factors became feasible.

To reduce false-positive clone identifications, we have assayed each pool with two different DNA probes in separate gelshift reactions. A DNA probe for the intended test was called the tester probe. The second probe, which we called the reference probe, formed a complex with a DNA-binding protein endogenous to the host cell. A pool was considered true positive when a protein-DNA complex was detected in the tester reaction without a gelshift intensity increase in the corresponding reference reaction. Intensity increases in both tester and reference reactions were interpreted as false-positive signals. By the comparison of tester and reference gelshift assays, we were able to distin-

guish between true- and false-positive pools and therefore increase the reliability of the clone selection.

MATERIALS AND METHODS

Division of the cDNA Expression Library

For expression cloning of mammalian brain transcription factors, a neonatal rat brain pEAK8 cDNA library was purchased (Edge BioSystems, Gaithersburg, MD, USA). An aliquot of the primary library, 500 000 cfu DH10B library host cells, was inoculated in 1 L Terrific Broth (Quality Biological, Gaithersburg, MD, USA) containing 100 mg/L ampicillin (Sigma, St. Louis, MO, USA). The library was divided into 960 pools, representing approximately 500 clones/pool complexity and further processed (Figure 1). Each pool in 1 mL volume was transferred into ten 96-well blocks supplied with Domed Lids to facilitate air convection (Edge BioSystems). The bacterial culture was incubated at 37°C for 16 h in a bacterial shaker-incubator at 320 rpm. After incubation, the A_{650} values of randomly tested samples reached 8 absorbance units.

Plasmid Preparation and Purification

Plasmid DNA was isolated from the pools using 96-Well Alkaline Lysis MiniPrep Kit, and the DNA samples were subsequently purified by the 96-Well Spin Columns/Gel Filtration Kit (Edge BioSystems). DNA concentrations were measured by SYBR® Green (Molecular Probes, Eugene, OR, USA) and by measuring fluorescence using DNA standards in a VICTOR² spectrofluorimeter (Perkin Elmer, Gaithersburg, MD, USA). For testing the clone selection technique by specific controls, before the library division, SATB1 (2) and HMGI(Y) (unpublished data) coding sequences were ligated into pEAK10 expression vectors (Edge BioSystems) and DH10B cells (Invitrogen, Carlsbad, CA, USA) were subsequently transformed with the constructs. The transformants were mixed with DH10B cells of randomly selected library pools in 1:50, 1:150, and 1:500 ratios and were processed for plasmid DNA production as described above. As an internal control, the SATB1 expression clone was mixed into the primary cDNA library in an estimated 1:500 000 cfu ratio respectively and was processed with the cDNA library.

Transfection of Mammalian Cells and Protein Expression

As a mammalian host for protein expression, embryonic kidney epithelial PEAK^{Rapid} Cells were purchased as part of the PEAK^{Rapid} Expression Kit (Edge BioSystems). Cells were plated on 24-Well Cell Culture Clusters (Corning, Acton, MA, USA) and transfected with the expression vector pools using the Calcium Phosphate Transfection protocol (Edge BioSystems) according to the manufacturer's instructions. After 48 h incubation, the culture media were aspirated from the wells, and the plates were placed on dry ice for 10 min. The plates were sealed and transferred to -80°C for further processing. Fifty microliters of cell lysis buffer [20 mM HEPES, pH 7.9, 450 mM KCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.1% IGEPAL, 25% glycerol, 0.5 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF),

Figure 1. Flow chart for the agarose thin-layer gelshift clone selection procedure.

Research Report

10 µg/mL leupeptin, and 10 µg/mL aprotinin (all from Sigma)] were added directly to the frozen cells. The plates were maintained on dry ice to allow the buffer to freeze above the cell monolayers. The plates were then moved to room temperature. Immediately after thawing the frozen solution, the plates were placed on ice and the cell monolayers were resuspended by pipetting until viscous fibers of genomic DNA were apparent, indicating the solubilization of nuclear content. The cell lysates were transferred to microcentrifuge tubes and/or 96-well microplates and stored at -80°C. Before the actual experiment, the transfection conditions were optimized by using the control SATB1 and HMGI(Y) expression vectors and by sampling protein production at different time points. From the test samples the DNA binding function of SATB1 and HMGI(Y) control proteins was monitored by agarose thin-layer gelshift assay.

Selection of Positive Pool by DNA-Binding Assay

Agarose thin-layers were prepared from 1% ULTRA™ pure agarose (Invitrogen) in 1× TBE buffer by bringing the suspension to a boil in a microwave oven. The clear melted agarose solution was moved to room temperature and allowed to cool to 65°C. A 200 × 200 mm agarose gel tray was adjusted to level position, and two Teflon combs were positioned sequentially in the tray. The gap between the edge of teeth and tray surface was approximately 1 mm. Derlin Combs (model H1/H4, 30-tooth, 2-mm; Invitrogen) were modified by trimming off one edge of the teeth at a 45° angle to focus the loaded sample. The distance between two teeth should match the dimensions of a multichannel pipettor that should be compatible with the wells of a microplate. Gel solution (180 mL) was poured into the tray, and the gel was allowed to solidify at room temperature in a level position. After solidification, the thickness of the gel was 2–2.5 mm. The gel was either covered with plastic wrap and stored at 4°C for a maximum of one week or used immediately. For the binding reaction, a tester probe (rat enkephalin gene -2496, -2556 sequences; GenBank® ac-

cession no. U03026) was chemically synthesized (Gene Probe Technologies, Gaithersburg, MD, USA) and radioactively labeled with 2500 cpm/fmol specific activity, annealed, and gel purified as described (3). The reference probe (rat enkephalin gene -496, -467) was prepared similarly. The binding reactions were carried out in pre-chilled, 60-well MicroWell Mini Trays (Nalge Nunc International, Naperville, IL, USA) or in 96-well microplates. Ten microliters of binding stock solution containing 12 mM HEPES, pH 7.9, 100 mM KCl, 10 mM NaCl, 1 mM MgCl₂, 0.5 mM AEBSEF, 10 µg/mL leupeptin, 10 µg/mL aprotinin, 5% glycerol, 0.1% BSA, 0.1 µg/µL poly[d(I-C)], and 1 fmol/µL radioactively labeled tester or reference probe were added to the wells of the pre-chilled plates. Cell lysates stored in microplates were thawed on ice, and 0.33 µL (approximately 0.7 µg total protein) of cell lysates were added to the reaction mixtures. The solutions were gently mixed using the pipet tips. The protein concentrations were adjusted before the actual experiment to maintain an excess of free probe in the reaction. The binding mixtures were incubated on ice for 5 min and then loaded into the wells of pre-chilled agarose thin-layers in 5-µL aliquots. Electrophoresis was carried out in 1× TBE buffer applying 200 V starting voltage (5 V/cm) in a horizontal gel electrophoresis unit (Stratagene, La Jolla, CA, USA) for 1 h at 4°C–8°C of controlled temperature to increase the caging effect that stabilizes protein-DNA complexes during electrophoretic gel migration (9). The gel was maintained in horizontal position during the electrophoresis, and the running buffer level was 2 mm above the gel surface. After the electrophoretic separation of bound and free fractions, the gel was carefully transferred onto quaternary amine derivatized Zeta-Probe GT Genomic Tested Blotting Membrane (Bio-Rad Laboratories, Hercules, CA, USA), placed on two layers of Whatman® 3MM Chr paper (Whatman, Clifton, NJ, USA), and covered with plastic wrap. Subsequent transfer was accomplished by simultaneous application of high-temperature (80°C) and less than 20 torr vacuum pressure for 25 min in a Gel Dryer model 583 (Bio-

Rad Laboratories) using a VLP 120 Vacuum Pump (Savant Instruments, Holbrook, NY, USA). The blotted gels were subjected to autoradiography and quantitative analysis using a Storm™ 860 PhosphorImager (Molecular Dynamics, Sunnyvale, CA, USA).

Subdivision of Positive Pool and Single Clone Selection

The identified single positive pool that represented only 500 different clones was further divided into 96 small pools. PEAK^{Rapid} cells were transfected with these small pools, and the expressed proteins were tested for DNA-binding as described in the DNA-binding assay section. A single small pool that represented five individual clones was sorted by PCR (Invitrogen). The individual clones were expressed in PEAK^{Rapid} cells and were tested for DNA-binding. The DNA sequence of the positive clone was determined (Biomedical Instrumentation Center, USUHS, Bethesda, MD, USA), and, for clone verification, polyclonal antibodies were obtained. For the supershift assay, 1 µL antibody was combined with 2 µg neonatal rat brain nuclear protein extract and tested in gelshift assay with the tester probe as described in the DNA-binding assay section. After 20 min incubation on ice, the reaction mixtures were analyzed by electrophoresis on a thin-layer agarose gel.

It is important to note that the protein yield from each expression pool is sufficient for the screening of the entire library with 100 different DNA probes.

RESULTS AND DISCUSSION

For cloning the cDNA of a mammalian DNA-binding protein, we have divided a high-complexity cDNA expression library into multiple low-complexity cDNA pools (Figure 1). Mammalian cells were subsequently transfected with the cDNA pools for protein expression. For clone selection, we employed an agarose thin-layer gelshift assay.

In a control experiment, we addressed the question of whether or not endogenous proteins, expressed by the host cell, would mask the gelshift of a

recombinant protein. HMGI(Y) protein is endogenous to dividing epithelial cells. Therefore, we co-expressed the HMGI(Y) expression vector with cDNA expression vector pools in various molar ratios. We could detect the control HMGI(Y) protein over the endogenous HMGI(Y) even in a low 1:500 ratio (Figure 2A). However, we cannot exclude the possibility that high levels of an endogenous DNA-binding protein might mask the gelshift of a positive clone. In a next set of experiments, we assessed how cDNA size-dependent bias might influence the library screening procedure. The cDNA size-depen-

dent clone representation bias may be due to disadvantageous transfection, transcription, translation, and post-translational modification circumstances that may ultimately decrease the representation of the larger size of proteins in expression pools. To assess this bias, HMGI(Y) protein (cDNA size, 1.6 kb) and SATB1 protein (cDNA size, 4.3 kb) were compared in gelshift assays. We found no indication of size-dependent bias in this control experiment (Figure 2A). However, disadvantageous plasmid replication during library vector pool amplification may also result in a cDNA size-dependent bias. To test this

Figure 2. Control proteins from cDNA expression pools detected by agarose thin-layer gelshift assay. (A) Expression vectors of HMGI(Y) or SATB1 proteins were mixed with cDNA expression vector pools in the indicated ratios. Protein pools were expressed in PEAKRapid cells and were tested with radioactively labeled tester probe in gelshift assay. The BCGR lane shows background-binding activity observed in gelshift assay with PEAKRapid cell extract. Bands of HMGI(Y) and SATB1 proteins are indicated by arrows. TESTER FP marks the free probe gel migration position. (B) Detection of SATB1 gelshift during the library screening procedure from expression pool 13 at row E in plate 7 (star). SATB1 expression clone was mixed into the primary cDNA library in an approximately 1:500,000 cfu ratio before the library division. The up-down arrow compares the binding intensities between tester and reference gelshift assays in the SATB1-containing pool. Migration positions of tester and reference free probes (FP) are marked by arrows.

Figure 3. Detection of a positive pool by agarose thin-layer gelshift and preliminary clone verification by antibody supershift assay. (A) A DNA-binding protein is detected in pool number 8 at row G in plate 6 by agarose thin-layer gelshift assay (star). The up-down arrow compares tester and reference gelshift reactions of the positive pool. Arrows indicate the gel migration position of tester and reference free probes (FP). (B) The DNA sequence encoding the positive binding activity was determined, and two specific antibodies (SPECIFIC ANTIB.) that recognize the protein were obtained. In supershift assays, the two specific antibodies or two control non-related antibodies (CONTROL ANTIB.) were added to neonatal brain protein extract (BRAIN EXTR.) and analyzed in supershift assay with the tester probe. Arrows indicate the migration positions of tester free probe (FP) and a gelshift band (DNA-PROTEIN CPX.) of a neonatal brain protein. Specific antibodies recognized this complex. Binding of specific antibodies to the complex resulted in supershifted bands (left brace).

Research Report

possibility, we mixed the SATB1 expression clone into the original primary cDNA library in an estimated 1:500 000 cfu ratio and processed the library for expression clone selection. As a result of library screening, we could identify the SATB1 binding activity in an expression pool (Figure 2B). Indeed, cDNA size-dependent bias did not interfere significantly with the clone selection procedure.

Because protein concentrations may vary between expression pools, it was difficult to interpret if an increased binding activity was due to the presence of a true-positive clone or if it was due to an occasionally higher endogenous protein concentration. To get around this difficulty, we tested each library pool with two different DNA probes in gelshift assays. When gelshift with the tester probe showed an increase in binding activity, we immediately examined the binding intensity of the corresponding reference gelshift reaction (Figure 2B). Variations in binding intensities in reference gelshift reactions directly indicated the variations in total protein concentrations. Therefore, gelshift reactions with reference probe provided a good measure to distinguish between true-positive and false-positive binding activity. Reference gelshifts in our experiments monitored the binding activity of an endogenous Oct-1-like element binding protein (3).

As an example of clone selection, we showed the detection of a positive pool (Figure 3A). Comparison of tester and reference gelshift intensities suggested that a true-positive pool was identified. We further divided this positive pool to select the sub-pool that contained the positive clone. Subsequent clone sorting resulted in a single cDNA clone. After determining the DNA sequence of the clone candidate, antibodies that recognize the encoded protein were obtained. As a preliminary step of clone verification, we tested if these antibodies could recognize the recombinant protein in a supershift assay (data not shown). Importantly, proteins from neonatal rat brain extract also tested positive in the same supershift assay (Figure 3B). This result was important because we prepared the original cDNA expression library from rat neonatal mRNA. Interestingly, we observed mul-

tiples supershifted bands in the supershift assay of neonatal rat brain extract that might indicate the presence of multiple protein isoforms or associated protein partners. An alternative approach of preliminary clone verification is the competitive gelshift assay to assess the specificity of protein-DNA interactions. However, the ultimate clone verification goal is the assessment of biological function. For the given example, the functional clone verification experiments are currently in progress.

In the clone selection technique described here, we showed that the combination of an expression cloning technology with agarose gelshift assay resulted in a sensitive clone selection technique. Moreover, to avoid false-positive clones, we have applied the tester-reference concept. Finally, it is conceivable that gelshift assays on agarose thin-layers may permit a high-throughput selection of factors that either facilitate or inhibit the binding activity of specific transcription factors.

ACKNOWLEDGMENTS

We thank the expert support of Dr. Reyes Candau (EdgeBio Systems). The SATB1 control reagent was provided by Dr. Terumi Kohwi-Shigematsu (Lawrence Livermore Laboratory, Berkeley, CA, USA). The research was supported by the Defense Veterans Head Injury Program. A.O. was supported by the USU Summer Student Program.

REFERENCES

1. **Cavalli-Sforza, L.L. and J. Lederberg.** 1955. Isolation of pre-adaptive mutants in bacteria by sib selection. *Genetics* 9:367-381.
2. **Dickinson, L.A., T. Joh, Y. Kohwi, and T. Kohwi-Shigematsu.** 1992. A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* 70:631-645.
3. **Dobi, A., M. Palkovits, C.G. Palkovits, M.A. Ring, and D.v. Agoston.** 2000. Septamer element-binding proteins in neuronal and glial differentiation. *J. Neurosci.* 20:1073-1084.
4. **Fried, M.G. and D.M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* 9:6505-6525.
5. **King, R.W., K.D. Lustig, P.T. Stukenberg, T.J. McGarry, and M.W. Kirschner.** 1997. Expression cloning in the test tube. *Science* 277:973-974.
6. **Li, J.J. and I. Herskowitz.** 1993. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science* 262:1870-1874.
7. **McCormick, M.** 1987. Sib selection. *Methods Enzymol.* 151:445-449.
8. **Tsai, S-F., D.I.K. Martin, L.I. Zon, A.D. D'Andrea, G.G. Wong, and S.H. Orkin.** 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* 339:446-451.
9. **Varshavsky, A.** 1987. Electrophoretic assay for DNA-binding proteins. *Methods Enzymol.* 151:551-565.

Received 19 December 2001; accepted 22 May 2002.

Address correspondence to:

Dr. Albert Dobi
Uniformed Services University
Department of Anatomy, Physiology,
and Genetics
4301 Jones Bridge Road, B2035
Bethesda MD 20814, USA
e-mail: adobi@usuhs.mil

**For reprints of this or
any other article, contact
Reprints@BioTechniques.com**