within autosomal genes are unavailable for assay validation. However, male samples are equivalent to a heterozygous deletion in female samples. Consequently, female samples with heterozygous deletions of an X chromosomal segment are comparable to normal male samples.

RD-PCR was also tested in a prospective study in hemophilia A families to detect heterozygous deletions in X chromosomal segments. RD-PCR assays were designed for exons 6 and 7 of the *factor VIII* gene. Eight females from two hemophilia A fam alidation of new assays; (*iv*) utilization of a generic endogenous internal dosage control to eliminate preparation and manipulation errors; (*v*) detection of gene dosage over a wide dynamic range (3); (*vi*) tolerance toward genomic DNA of variable quality (J. Shi, Q. Liu, V.Q. Nguyen, and S.S. Sommer, submitted); and (*vii*) uniform and unbiased performance across regions of variable sequence context and GC content.

ACKNOWLEDGMENTS

The authors would like to thank Xuemin Li for preparation of the blinded DNA samples.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing interests.

REFERENCES

1. **Stenson, P.D., E.V. Ball, M. Mort, A.D. Phillips, J.A. Shiel, N.S. Thomas, S. Abeysinghe, M. Krawczak, et al.** 2003. Human Gene Mutation Database (HGMD): 2003 update. Hum. Mutat. *21*:577-581.
2. **Armour, J.A., D.E. Barton, D.J. Cockburn, and G.R. Taylor.** 2002. The detection of large deletions or duplications in genomic DNA. Hum. Mutat. *20*:325-337.
3. **Liu, Q., X. Li, J.S. Chen, and S.S. Sommer.** 2003. Robust dosage-PCR for detection of heterozygous chromosomal deletions. BioTechniques *34*:558-570.
4. **Liu, Q. and S.S. Sommer.** 1998. Subcycling-PCR for multiplex long-distance amplification of regions with high and low GC content: application to the inversion hotspot in the factor VIII gene. BioTechniques *25*:1022-1028.
5. **Miyada, C.G. and R.B. Wallace.** 1987. Oligonucleotide hybridization techniques. Methods Enzymol. *154*:94-107.
6. **Yoshitake, S., B.G. Schach, D.C. Foster, E.W. Davie, and K. Kurachi.** 1985. Nucleotide sequence of the gene for human factor IX (anti-hemophilic factor B). Biochemistry *24*:3736-3750.

Received 6 April 2004; accepted 28 May 2004.

Address correspondence to Steve S. Sommer, Chair, Departments of Molecular Genetics and Molecular Diagnosis, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, CA 91010-3000, USA. e-mail: sommerlab@coh.org

Scalable high-throughput micro-expression device for recombinant proteins

Rebecca Page, Kin Moy, Eric C. Sims, Jeffrey Velasquez, Brendan McManus, Carina Grittini, Thomas L. Clayton, and Raymond C. Stevens

Joint Center for Structural Genomics and Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA, USA

BioTechniques 37:364-370 (September 2004)

The large-scale expression and purification of recombinant proteins needed for structural studies is time-consuming and expensive, especially when costly reagents, such as selenomethionine (SeMet) or $^{15}N/^{13}C$-labeled nuclear magnetic resonance (NMR) medium, are required. Small-scale screens are being developed to identify which targets express and are soluble prior to large-scale expression (1–7). Although these screens are designed for a 96-sample, small-scale format, none adequately predicts the reliable expression behavior with scaled-up milliliter and liter fermentations, which has been a challenge for researchers in the field.

We have adapted a low-cost, high-velocity incubating Glas-Col Vertiga shaker (Glas-Col, LLC, Terre Haute, IN, USA; Figure 1A) to develop an efficient, high-throughput *Escherichia coli* microliter-scale expression screening protocol that accurately predicts protein behavior expressed in large-scale (milliliter and liter) fermentation conditions. The apparatus shakes cultures in three dimensions at speeds of up to 1000 rpm, allowing small-scale (approximately 500 µL) cultures grown in 2-mL 96-deep-well blocks to achieve absorbances (A_{600}) as high as 10–20. This generates sufficient material for the analysis of expression, solubility, and binding to affinity purification matrices. Moreover, this screening strategy has also been used to identify clones that express and are soluble under SeMet (8) or $^{15}N/^{13}C$-labeled expression conditions that are necessary for the production of labeled recombinant proteins for direct structural analysis. The Glas-Col Vertiga shaker can be used to screen for soluble expression clones under native, SeMet, and $^{15}N/^{13}C$-labeled expression conditions.

Microscale expression using the Glas-Col Vertiga shaker is carried out as follows. First, overnight cultures are prepared with 250 µL sterile 2xyt media containing antibiotic (0.3 mM ampicillin) in sterile 2-mL, 96-deep-well, round-bottom blocks (USA Scientific, Ocala, FL, USA). Each well is then inoculated with 1 µL glycerol stocks of expression clones (also stored in a 96-well format) using a multichannel pipet. Once sealed with a sterile, porous plate cover (USA Scientific), the blocks are placed in the Glas-Col Vertiga shaker, and cultures are grown overnight at 37°C, shaking at 550 rpm.

Expression under native conditions is carried out using 500 µL sterile Terrific Broth (TB or 2× TB) media prepared with 100 mM MOPS, pH 7.4, and antibiotic, while expression under SeMet conditions is carried out using sterile complete media with 1 µg/mL $FeSO_4$, 0.8 mM SeMet, and antibiotic. Blocks containing the overnight cultures are then centrifuged at 5000× *g* to pellet the *E. coli* cells, the supernatants are discarded, and the pellets are resuspended in 100 µL fresh media (TB or SeMet, as appropriate). The freshly prepared expression blocks are then inoculated with 25 µL of the resuspended overnight culture. After sealing each expression block with a porous plate cover, blocks are placed in the Glas-Col Vertiga shaker and grown at 900 rpm and 37°C. When the culture A_{600} reaches approximately 3 (about 1.5 h following inoculation at 37°C; spectrophotometer wavelength, 600 nm), expression is induced [final concentration of 0.2% (w/v) L-arabinose (Sigma, St. Louis, MO, USA)] and cultures are allowed to shake at 900 rpm and 37°C for an additional 4–5 h. The final average culture A_{600} is usually between 10 and 20. Blocks are then centrifuged at 5000× *g* for 10 min to pellet the cells, and the supernatants are discarded. Blocks with pelleted cells are then analyzed for expression or sealed with a rubber plate seal and stored at -80°C for future analysis. In addition to the L-arabinose expression system (pBAD vector; Invitrogen, Carlsbad, CA, USA), similar microliter-scale and milliliter/liter experiments have also been carried out with the T7 expression system (pET vectors; Novagen, Madison, WI, USA).

Expression and solubility are evaluated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot analysis (His-tag antibodies). First, frozen bacterial cell pellets are placed at room temperature and allowed to thaw slightly. Cells are then thoroughly resuspended in 100 µL of MP lysis buffer A [50 mM Tris, pH 7.5, 50 mM sucrose, 1 mM EDTA, 0.25 mM tris-(2-carboxyethyl) phosphine hydrochloride (TCEP)] with 1.0 mg/mL lysozyme and 0.25 µL/mL Benzonase® (Novagen). After incubating at room temperature for 15 min, 100 µL of MP lysis buffer B (10 mM Tris, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 1 mM EDTA, 0.25 mM TCEP) are added to each well and incubated for an additional 15 min. After mixing, 10 µL samples of the lysate from each well are removed and added to the SDS-PAGE sample buffer. These samples represent the "total lysate" fractions. Samples are then centrifuged at 6100× *g* for 15 min to pellet cell debris. Following centrifugation, 20 µL of the supernatant are removed for SDS-PAGE analysis. These samples represent the "soluble" fractions.

If desired, the soluble fraction is then purified using immobilized metal-affinity chromatography (IMAC). A 400-µL, 0.45-µm filter-containing 96-well microplate is used for the load, wash, and elution steps, and a 450-µL 96-well plate is placed underneath the filter plate to collect the flow through. One hundred microliters of a 50% slurry of Co^{2+}-charged metal-affinity resin (TALON™ Superflow; BD Biosciences Clontech, Palo Alto, CA, USA) are added to each well of the filter plate and equilibrated with 300 µL of equilibration buffer (1:1 mixture of MP lysis buffers A and B). Equilibration buffer is removed by centrifugation at 300× *g* for 1 min. Clarified lysate supernatants are then added to each well and gently mixed by slowly pipetting up and down for 10 repetitions. The resin is allowed to settle to the bottom of each well and is then centrifuged as before, with new collection plates underneath to collect soluble lysate flow-through fractions. The resin is then washed once with MP wash buffer (25 mM Tris, pH 7.8, 150 mM NaCl, 5 mM imidazole, 0.25 mM TCEP) and centrifuged as before, saving the flow-through fractions in a third plate. The bound proteins are eluted from the IMAC resin using MP elution buffer (25 mM Tris, pH 7.8, 150 mM NaCl, 150 mM imidazole, 0.25 mM TCEP). The resin is slowly resuspended with 50 µL of elution buffer, after which the resin beads are allowed to settle for 10 min. The filter plates are then centrifuged a final time, using a new 450-µL 96-well plate to collect the eluate. Purified proteins can be used for additional analytical studies, such as mass spectrometry or size exclusion chromatography, as desired. A complete analysis of recombinant protein expression is performed by running 4%–20% SDS-PAGE gels of the total lysate, soluble, flow-through (unbound), and eluate fractions.

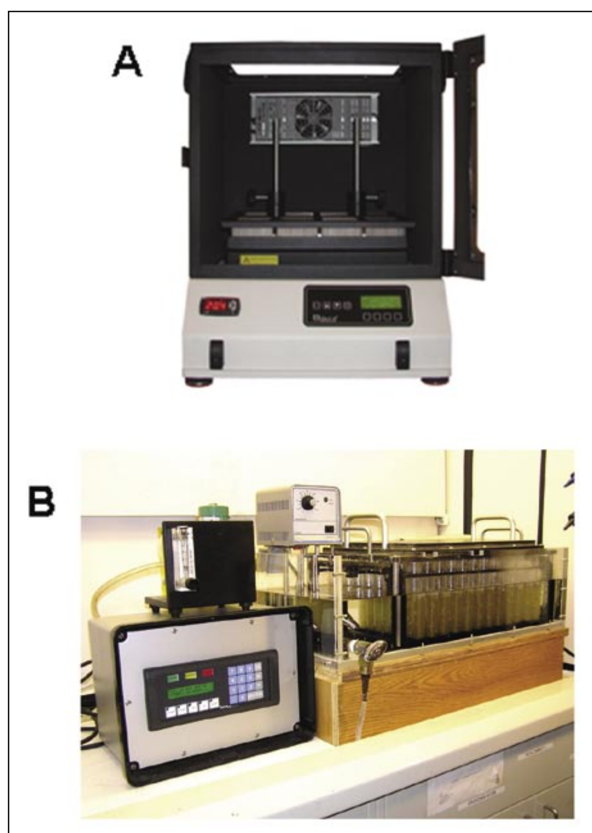

Figure 1. Glas-Col Vertiga microliter and Genomics Institute of the Novartis Research Foundation (GNF) milliliter expression systems. (A) Glas-Col Vertiga microliter-scale expression system. A low-cost vertical shaker apparatus adapted for the development of a scalable high-throughput *Escherichia coli* micro-expression screening. (B) GNF 96-well milliliter fermenter (9). A typical run includes the growth of approximately 65 mL of growth media, grown to an absorbance (A_{600}) of 10–20 using media conditions similar to the microliter-scale expression experiments, with oxygen supplementation.

Table 1. Comparison List of Microliter-Expressed and Scaled-Up Expressed Proteins

GenBank Accession No.	Protein Name	Macro-Soluble	Micro-Soluble
2650105	Chloroplast inner envelope membrane protein (*Archaeoglobus fulgidus*)	S	S
2648890	Ornithine cyclodeaminase (arcB) (*A. fulgidus*)	S	S
2648361	Conserved hypothetical protein (*A. fulgidus*)	S	S
2648357	Conserved hypothetical protein (*A. fulgidus*)	S	S
2649029	*A. fulgidus* predicted coding region AF1543	S	S
2649912	Conserved hypothetical protein (*A. fulgidus*)	S[a]	S[a]
2649331	Oxaloacetate decarboxylase, subunit alpha (oadA) (*A. fulgidus*)	S	S
2650039	Conserved hypothetical protein (*A. fulgidus*)	S	S
2650115	*A. fulgidus* predicted coding region AF0514	S	S
15162326	AGR_pAT_752p (*Agrobacterium tumefaciens*)	S	S
15159470	AGR_L_2016p (*A. tumefaciens*)	S	S
15156516	AGR_C_2657p (*A. tumefaciens*)	S	S
15159614	AGR_L_2275p (*A. tumefaciens*)	S	S
15155223	Ubiquinone/menaquinone biosynthesis methyltransferase AGR_C_559p (*A. tumefaciens*)	S	S
10172709	Transcriptional regulator involved in nitrogen regulation (*Bacillus halodurans*)	S	S
10173683	Beta-xylosidase (*B. halodurans*)	S	S
10176051	p-nitrophenyl phosphatase (*B. halodurans*)	S	S
10175117	L-serine dehydratase beta subunit (*B. halodurans*)	S	S
10174212	BH1595: unknown conserved protein (*B. halodurans*)	S	S
10174951	BH2331: unknown conserved protein (*B. halodurans*)	S	S
10173897	BH1281: unknown conserved protein in others (*B. halodurans*)	S	S
2636534	Alternate gene name: yxa Q (*Bacillus subtilis*)	S	S
AAD35532	Thymidylate synthase thyX (*Thermotoga maritima*)	S	S
2649518	Coenzyme F390 synthetase (ftsA-1) (*A. fulgidus*)	S	S
10173275	Response regulator aspartate phosphatase (*B. halodurans*)	S	S
10173447	BH0832: unknown conserved protein in others (*B. halodurans*)	N, S[b]	S, S[b]
10175361	Proline oxidase (proline dehydrogenase) (*B. halodurans*)	N	N
15160206	AGR_L_3243p (*A. tumefaciens*)	N	N
15159665	Antioxidant enzyme AGR_L_2357p (*A. tumefaciens*)	N, S[b]	N, N[b]
2650313	Rhamnosyl transferase (rfbQ) (*A. fulgidus*)	N	N
2650194	3-ketoacyl-CoA thiolase (acaB-8) (*A. fulgidus*)	N	N
2650329	*A. fulgidus* predicted coding region AF0307	N	N
10173104	BH0491: unknown conserved protein in others (*B. halodurans*)	N	N
10174332	Thioredoxin (*B. halodurans*)	N	N

The experiments were carried out in duplicate. If expression results were not identical between experiments, the results from both were reported as S, soluble, or N, insoluble or no expression. Thirty-two of thirty-four proteins were consistently expressed or not expressed in both sets of experiments. The expression results that were not correlated for proteins BH0832 (10173447) and AGR_L_2357p (15159665) are highlighted, and in both cases, the lack of correlation was observed in only one of the two experiments.
[a]This protein failed to express in experiment 2.
[b]Expression results from two separate sets of experiments. If only a single letter is shown, then results were identical between the two experiments.

For comparison to larger-scale fermentation behavior, 34 proteins were expressed using the microliter-scale expression protocol under native conditions and expressed again using the Genomics Institute of the Novartis Research Foundation (GNF; San Diego, CA, USA) 96-well milliliter fermenter (Figure 1B). This experiment was carried out in duplicate. Targets expressed in the GNF 96-well fermenter were purified as outlined by Lesley, et al. (9), with purity and yield determined by SDS-PAGE. Most samples were further analyzed using Western blot analysis with an anti-His antibody (Sigma) using standard protocols. Of the 34 proteins tested, 32 consistently expressed (or did not express) in both microliter and larger-scale volumes, illustrating a high level of scalability and correlation in

the expression levels of soluble protein between the Vertiga screening trials and large-scale growth conditions (Table 1). In contrast, without the Vertiga shaker, the correlation between microliter- and milliliter- (and liter) scale expression is much lower (data not shown). The two proteins that behaved differently between micro- and macro-expression, BH0832 (GenBank® accession no. 10173447) and AGR_L_2357p (accession no. 15159665), did so in only one of the two experiments; in the second, the micro- and macro-expression behavior was identical for both proteins (Table 1). Finally, eukaryotic proteins from *Saccharomyces cerevisiae* and viral proteins from the Severe Acute Respiratory Syndrome (SARS) virus have also been successfully micro-expressed in *E. coli* using this device.

These results demonstrate that the microliter-scale expression protocol developed using the Vertiga vertical shaker can be used to accurately identify proteins that will express solubly in larger-scale fermentation conditions. Moreover, the results from these screens can be used to assess the solubility and expected protein yield for each potential protein target for both native and labeled *E. coli* recombinant expression. Future developments include using material directly from microliter-scale expression for biophysical analysis, including nanovolume crystallization (10) and other biophysical techniques currently being miniaturized.

ACKNOWLEDGMENTS

This work was supported in part by grant no. GM62411 (to I.A. Wilson, P.I., JCSG) from the National Institutes of Health (NIH) Protein Structure Initiative. We appreciate the assistance of Lee Clark, Jim Jasco from Glas-Col, Mark Knuth and Ciaran Cronin from Syrrx for the initial studies of the Vertiga system, and Scott Lesley, Mark Weselak, and Bob Downs for the GNF fermenter design.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing interests.

REFERENCES

1. **Holz, C., O. Hesse, N. Bolotina, U. Stahl, and C. Lang.** 2002. A micro-scale process for high-throughput expression of cDNAs in the yeast *Saccharomyces cerevisiae*. Protein Expr. Purif. *25*:372-378.
2. **Shih, Y.-P., W.-M. Kung, J.-C. Chen, C.-H. Yeh, A.H.-J. Wang, and T.-F. Wang.** 2002. High-throughput screening of soluble recombinant proteins. Protein Sci. *11*:1714-1719.
3. **Adams, M.W.W., H.A. Dailey, L.J. DeLucas, M. Luo, J.H. Prestegard, J.P. Rose, and B.-C. Wang.** 2003. The Southeast Collaboratory for Structural Genomics: a high-throughput gene to structure factory. Acc. Chem. Res. *36*:191-198.
4. **Nguyen, H., B. Martinez, N. Oganesyan, and K. Rosalind.** 2004. An automated small-scale protein expression and purification screening provides beneficial information for protein production. J. Struct. Funct. Genomics *5*:23-27.
5. **Scheich, C., V. Sievert, and K. Büssow.** 2003. An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography. BMC Biotechnol. *3*:12-19.
6. **Knaust, R.K.C. and P. Nordlund.** 2001. Screening for soluble expression of recombinant proteins in a 96-well format. Anal. Biochem. *297*:79-85.
7. **Yokoyama, S.** 2003. Protein expression systems for structural genomics and proteomics. Curr. Opin. Chem. Biol. *7*:39-43.
8. **Hendrickson, W.A., J.R. Horton, and D.M. LeMaster.** 1990. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J. *9*:1665-1672.
9. **Lesley, S.A., P. Kuhn, A. Godzik, A.M. Deacon, I. Mathews, A. Kreusch, G. Spraggon, H.E. Klock, et al.** 2002. Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. Proc. Natl. Acad. Sci. USA *99*:11664-11669.
10. **Santarsiero, B.D., D.T. Yegian, C.C. Lee, G. Spraggon, J. Gu, D. Scheibe, D.C. Uber, E.W. Cornell, et al.** 2002. An approach to rapid protein crystallization using nanodroplets. J. Appl. Crystallogr. *35*:278-281.

Received 19 April 2004; accepted 1 June 2004.

Address correspondence to Raymond C. Stevens, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. e-mail: stevens@scripps.edu

Robust expression of transgenes in MCF-7 breast cancer cells is expression vector-dependent

Paul Winnard, Jr., Yelena Mironchik, and Venu Raman

The Johns Hopkins University, Baltimore, MD, USA

BioTechniques 37:370-374 (September 2004)

We are reporting that the in vivo protein expression in mammalian MCF-7 breast cancer cells can be highly dependent on the type of commercially available expression vectors used. Figure 1 is a representation of the vector backbones used and indicates the position of our genes of interest (denoted Insert) relative to the cytomegalovirus (CMV) promoter and simian virus 40 (SV40) poly(A) signal sequence in a pCMVTag4A vector (Stratagene, La Jolla, CA, USA), which resulted in an in-frame fusion at the 3′ end with a FLAG® epitope sequence. Figure 1 also depicts the placement of these FLAG-tagged genes (including a TGA stop codon) into a pEF-1α-His/Myc vector (Invitrogen, Carlsbad, CA, USA) and their relative positions with respect to the elongation factor-1α (EF-1α) promoter and bovine growth hormone (BGH) poly(A). Thus, except for the vector backbones, both fusion gene construct sequences were identical in the two vector systems. Sequencing analyses proved that each gene sequence was correct and in-frame with the FLAG sequence and stop codon. All plasmids were prepared using the Qiagen® Plasmid Midi Kits (Qiagen, Valencia, CA, USA). Initial transient