Method to express and purify nm23-H2 protein from baculovirus-infected cells

L. Garzia¹, A. André¹, A. Amoresano², A. D'Angelo¹, R. Martusciello¹, C. Cirulli², T. Tsurumi³, G. Marino², and M. Zollo¹

¹Telethon Institute of Genetics and Medicine, Naples, Italy, ²School of Biotechnological Sciences, Università di Napoli Federico II, Naples, Italy, and ³Aichi Cancer Center Research Institute Division of Virology, Nagoya, Japan

BioTechniques 35:384-391 (August 2003)

High-throughput protein expression and purification are major bottlenecks in the postgenomic and proteomic era. We show here an automated method to express and purify nm23-H2, a nucleoside diphosphate kinase (NDPK), in a 96-well format, by the use of a robotic workstation, from insect Spodoptera frugiperda (Sf9) baculovirus-infected cells using nickel-nitrilotriacetic acid (Ni-NTA) agarose beads. The automated method is coupled to mass spectrometry for a validation and quality-control analysis. To verify the bona fide of the recombinant protein, several tests have been produced, including NDPK assay, Western blotting, and in vitro phosphorylation experiments, thus confirming the value of the protocol developed. The method has been validated for the expression of several proteins, thus confirming the value of this automated protocol. The research presented here is a useful method both for industrial and academic environments to produce in a high-throughput mode recombinant eukaryotic proteins to be assayed for a specific function in a systematic manner.

INTRODUCTION

With the evolution of genome projects and the improvements in DNA sequencing technology, several organisms sequenced have been deposited in public domain databases [i.e., Archaea, bacteria, eukaryotes, viruses, and organelles; see the National Center for Biotechnology Information (NCBI) Web site at <http://www.ncbi.nlm.nih.gov/>]. The vast number of proteins predicted and/or verified by comparing with a cDNA database is creating massive opportunities for both basic and applied research. One of the major efforts in the postgenomic and functional genomic era is to understand how unknown gene products cooperate in cellular environments to exert their cellular function. Efficient and rapid expression of genes in homologous/heterologous protein expression systems and rapid purification steps are today the major bottlenecks.

To date, methods have been generated to express simultaneously a high number of proteins, based on *Escherichia coli* and *Saccharomyces cerevisiae* expression systems. Both of these systems are adapted to grow and express proteins, in a 96-well format, to obtain

partially or fully automated processes. The baculovirus expression system has not yet been tested for its usefulness to express protein in this format, mainly because of its low efficiency on a 96-well-format growth and its incapacity to obtain a reasonable amount of purified protein. Nevertheless, a high yield in a single preparative heterologous expression system has been reported to date. As far as the novel eukaryotic gene expression systems are concerned, there are many advantages in using baculovirus for heterologous gene expression (1). As with other eukaryotic expression systems, baculovirus expression of heterologous genes permits folding, posttranslational modification, and oligomerization in manners that are close, often identical, to those that occur in mammalian cells. Low levels of expression can often be increased with the optimization of time of expression and multiplicity of infection (MOI). Very recently, the use of molecular chaperones (2), ready cloned in baculovirus vectors, has completely encompassed the potential misfolding and solubility problems in insect cells (3).

To investigate if this system should be used in experiments that include the

semi-automated expression of a high number of proteins, we evaluated the yield and the quality of the protein expression and purification procedure carried out in a 96-well format.

The recombinant baculovirus used in this work encoded for nm23-H2, a protein for which a nucleoside diphosphate kinase (NDPK) and histidine phosphotransferase characterized activity, a potential casein kinase II phosphorylation site (4), and a DNA binding and a c-Myc transactivating activity were found in vitro and in vivo (5–7).

Today there is an increasing demand for automated protein purification because, to date, the above step is widely considered to be a bottleneck in the assembly line of recombinant protein production. Automation would actually allow one to avoid repeating column chromatographic steps and, as an added benefit, also allows the standardization of the amount of protein produced. This issue has been focused on the *E. coli* expression environment, suggesting the use of nickel-nitrilotriacetic acid (Ni-NTA)-coated multi-well microplates (8). The present state of the art in our laboratory is based on the use of Ni-NTA magnetic agarose beads and their affinity to 6×His-tag protein directly isolated from the lysis buffer environment as presented by Lanio et al. (8). Compared to this method, the advantage of using agarose magnetic beads versus coated plates is the overall increased protein maximum binding capacity of the beads (0.2 nmol agarose Ni-NTA beads/well vs. 10 pmol coated Ni-NTA plates/well), thus resulting in a significant reduction in the purification costs. A second advantage is that the method allows host growing and several purification steps in the same multi-well plate. Finally, the magnetic agarose Ni-NTA beads can be used at least twice without affecting the protein purification rate.

Automation is not meant to obtain large quantities of protein but rather to produce automated simultaneous purification processes for several samples, such as purifying proteins in a 96-well plate simultaneously (a useful method for protein mutational analysis). The following are further advantages of using such an approach: (i) the sonication steps, often required to cleave genomic DNA, can be avoided; (ii) the fishing of

6× His proteins from cell lysates can be iterated many times to saturate the beads, thus improving their loading capacity; (iii) the method has been implemented with magnetic robotic stations. Indeed, the latter devices have already been introduced in the case of DNA purification in the assessment of the human genome draft sequence and can eventually be implemented in automating the protein purification, too.

Appended to the purification step, a quality-control analysis of recombinant products has been implemented by the use of direct analysis of protein-loaded beads using matrix-assisted laser desorption ionization (MALDI) mass spectrometry. Indeed, mass spectrometry plays an important role in proteomics, particularly in protein identification analysis, but also in the characterization of posttranslational modifications. In conjunction with classical protein and carbohydrate chemistry, these strategies have proved to be powerful and cost-effective methods for structural analysis, thus playing a key role in the quality control of recombinant proteins and glycoproteins (9–11). The first step in the elucidation of protein structure generally consists in the determination of the accurate molecular weight of the intact protein. The accuracy of the molecular mass determination confirms, or otherwise the encoded protein sequence suggests, the presence of unexpected covalent modifications. The nature and fine location of any structural modification in the protein sequence can then be identified and assigned using the mass-mapping strategy. Errors of translation, deletion, insertion, point mutation (except inversions in the sequence), posttranslational modifications or processing, and the S-S bridge pattern can be detected and assigned using this method. The sites and the nature of modifying groups originated by post-biosynthetic events (disulfide bonds, methylation, phosphorylation, hydroxylation, or glycosylation) can be precisely assessed. Some strategies have been specifically developed to characterize both N- and O-linked glycans, which integrate chemical manipulations, enzymatic degradation, and mass spectrometry analytical techniques. The definition of the entire structure of the oligosaccharide moieties released from glycoproteins is achieved by

determining the microheterogeneity of the different structures, the composition of the monosaccharides present, the branching pattern, the sequence of the antennae, and the possible occurrence of modifying groups, such as sulfate, phosphate, acetyl groups, etc. (12).

MATERIALS AND METHODS

Cell Culture and Baculovirus Stock Preparation

Spodoptera frugiperda (Sf9) cells were cultured in TNM-FH media (Sigma, Milan, Italy) with 10% fetal bovine serum (FBS), 100 U/mL penicillin, and 100 µg/mL streptomycin at 27°C. The cDNA for nm23-H2 was cloned in pFastBacHTa (Invitrogen, Carlsbad, CA, USA), and the recombinant baculovirus was obtained according to the manufacturer's instructions. With a cycle of amplification in the Sf9 cell at an MOI of 0.1, we obtained a virus stock preparation with an approximate titer of 2×10^7 pfu/mL.

Automated Protocol for Protein Expression and Purification

Sf9 cells were plated at a density of 5×10^5 /mL in a 96-well flat-bottom plate (Qiagen, Hilden, Germany) and infected at an MOI of 2 and 5 for 48 h at 27°C. The recombinant protein was first purified by Ni-NTA agarose magnetic beads (Qiagen) using 20 µL 5% suspension directly added to the crude lysate, manually applying the method created for the liquid handling workstation. Protein expression and purification under this condition were verified by Western blot analysis following a standard protocol. For the execution of the semi-automated method with a liquid handling workstation, Sf9 cells were grown using the condition previously optimized for the culture in the 96-well format. The infection was performed at an MOI of 2. After 48 h, the cells were treated by the liquid handling workstation (Multi-PROBE® II HT-EX; Packard Instruments; PerkinElmer Life Sciences, Boston, MA, USA), the supernatant (150 µL) was removed, and the cell was lysed by 100 µL lysis buffer (50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 10

mM imidazole, 0.005% Tween® 20) by up-and-down pipetting with conductive disposable tips (Molecular BioProducts, San Diego, CA, USA). The lysates were then supplemented with 20 µL 5% suspension of Ni-NTA agarose magnetic beads and incubated at 10°C on an Eppendorf® Thermomixer Comfort (Eppendorf, Hamburg, Germany) at 500 rpm for 2 h. The 96-well plate was then placed on the type-A magnet (Qiagen), and the flow-through was then aspirated from the lysis plate with conductive disposable tips and pipetted in a replica plate, which was stored for successive analyses. The lysis plate was removed from the magnet, and the beads were resuspended and washed with 150 µL wash buffer (50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.005% Tween 20) by up-and-down pipetting using the syringe-mode configuration. This washing step was repeated twice; in the last step, 150 µL of the bead suspension were transferred in a round-bottom 96-well plate to increase protein recovery. Then, the recombinant protein was eluted, incubating the beads in 50 µL elution buffer (50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 250 mM imidazole, 0.005% Tween 20) on the shaker for 5 min at 10°C. The round-bottom plate was then positioned on the magnet tool for 5 min, and the eluates were collected in a replica plate for further analysis on the thermomixer block. Accessory material describing detailed robotic settings and scripts ready to run on the Packard robot are available at (<http://seqcore.tigem.it/baculo>).

Protein Quantification

The recombinant protein obtained was quantified according to the Bradford method directly in the replica plate obtained from the purification step, using a 96-well ELISA microplate reader (model 550; Bio-Rad Laboratories, Hercules, CA, USA) with bovine serum albumin (BSA) as standard.

Western Blot Analysis

Elution fraction (35 µL) obtained from the protein purification protocol was loaded on 12.5% (w/v) polyacrylamide gel for sodium dodecyl sulfate polyacrylamide gel electrophoresis

(SDS-PAGE) and electroblotted onto a polyvinylidene difluoride (PVDF) membrane (Immobilon™-P; Millipore, Bedford, MA, USA). Monoclonal antibody penta-HIS (Qiagen) anti-HIS epitope was used according to the manufacturer's protocol to detect insect-expressed proteins. The proteins were further analyzed by a polyclonal antibody (Novocastra Laboratories, Peterborough, UK) that recognizes nm23-H2; this antibody was used at 1:500 dilution.

NDPK Assay

We used the method described previously (13), using 100 ng chromatography resin-purified proteins nm23-H1 and nm23-H2. The control killing activity mutation in the nm23-H2H118F protein was made as described previously (14). The slope of a linear plot of absorbance versus time was used as a measure of enzymatic activity.

In Vitro Protein Phosphorylation

Intact protein immobilized on mag-

netic beads was incubated in 1× reaction buffer (50 µL) supplemented with 200 µM ATP for 1 h at 30°C. Two units of recombinant casein kinase I (New England Biolabs, Beverly, MA, USA) were used. Casein kinase I was then eliminated with three washes with double-distilled water; the beads with the intact phosphorylated protein were resuspended in double-distilled water and immediately analyzed by mass spectrometry.

Mass Spectrometry Analysis

MALDI mass spectrometry analyses were carried out using a linear Voyager™ DE or a reflectron Voyager DE-PRO mass spectrometer (Applied Biosystems, Foster City, CA, USA) equipped with the delayed extraction device. Mass calibration was performed with apomyoglobin at 16951.5 Da and BSA at 66431.4 Da as internal standards for intact protein and with bovine insulin (average molecular mass 5734.6 Da) and a matrix peak (379.1 Da) for the tryptic mixtures. Ten microliters of 0.1% trifluoroacetic acid (TFA) were

Figure 1. Method describing growth, infection, and purification of nm23-H2 protein.

added to the beads, and 1 μ L bead suspension was mixed with 1 μ L sinapinic acid (25 mg/mL) in $\text{CH}_3\text{CN}/0.2\%$ TFA 70:30 (v/v) and air-dried. Fifty microliters of 50 mM ammonium bicarbonate, pH 8.5, were directly added to the bead suspension. Trypsin hydrolysis of the protein was carried out at 37°C for 18 h using 1:50 (w/w) enzyme-to-substrate ratio. Tryptic mixtures were dissolved into a final concentration of 10 μ M, and 1 μ L was applied to a sample slide and mixed with 10 mg/mL α -cyano-4-hydroxycinnamic acid solution in $\text{CH}_3\text{CN}/0.2\%$ TFA 70:30 (v/v) before air-drying. Mass spectra were generated from the sum of 500 laser shots.

RESULTS AND DISCUSSION

We at first defined the best MOI to infect the Sf9. At an MOI of 2, the protein production reached the plateau as confirmed by Western blot analysis (see Figure 2). The same conditions were applied to the use of chaperone proteins described previously (10). Co-infection with baculovirus chaperons Hsp40 and Hsp70 increased protein folding, resulting in 50 ng protein/well (data not shown). We used 20 μ L magnetic bead suspension; this amount has a binding capacity of 0.2 nmol, largely exceeding the quantity of the recombinant protein expected from previous

preparative experiments. The semi-automated purification procedure was performed from a 96-well plate infected for 48 h at an MOI of 2.

After the identification of the right MOI infection conditions, we have infected 96-well plate containing 5×10^5 seed Sf9 cells in TNM-FH insect cell-supplemented media. After 48 h incubation at 27°C, we have isolated the multi-well and performed the purification on the robotic workstation (Figure 1). The method and conditions totaled 3 h. The purified protein product for each well was measured by a colorimetric assay and evaluated as 50 ng protein purified/well, totaling almost 5 μ g purified

Figure 2. Validation of the protein purification protocol. (A) Identification of the correct multiplicity of infection (MOI) of Sf9 cells, ranging from 2×10^5 to 1.2×10^6 total number of cells, using two different virus concentrations (MOI of 2 and 5). (B) Protein quantity determination performed on randomly selected wells, from the 96-well plate acquired by the method. (C) Western blot analysis using polyclonal antibodies (anti-nm23-H2) used at 1:500 dilution. As positive control, 50 ng nm23-H2 purified protein were used. (D) Nucleoside diphosphate kinase (NDPK) assay to determine the specific activity of recombinant protein obtained from the method. As a negative control, a killer catalytic site mutant (nm23-H118F) was used. The y-axis shows U/mg protein identified. (E) Protein spectrum by MALDI analyses directly obtained from recombinant protein immobilized on magnetic agarose nickel-nitrilotriacetic acid beads.

fied protein/96-well plate.

Following the steps presented in Figure 1, the protein preparation was checked for activity (NDPK assay), identity (Western blot analysis and MALDI mass spectrometry), and post-translational modifications (tryptic digestion and MALDI mass spectrometry). Through Western blot analysis

(Figure 2, panels A and C), we did not observe any product derived from an incomplete protein synthesis or any degradation products.

Because the protein produced is a nucleoside diphosphate kinase, we applied the assay to verify its activity. The NDPK assay applied is the same presented in Reference 14. The nm23-

Figure 3. In vitro phosphorylation and MALDI mass spectrometry analyses of nm23-H2 purified protein. (A) A graphic representation of His-nm23-H2 protein attached to agarose nickel-nitrilotriacetic acid beads and a casein kinase I (CKI) protein motif recognition site of phosphorylation on nm23-H2 protein. (B) In red, protein sequence obtained by tryptic digestion resolved by MALDI; white sequences correspond to regions not solved. In blue, serines 120, 122, and 125 found phosphorylated. (C) Partial MALDI mass spectrometry spectrum of the tryptic digest nm23-H2 phosphoprotein. The phosphorylated peptides at m/z 1069.6 and 1151.5 (1), 1232.6 (2), and 515.3 (3) are reported in the panels.

H2 protein produced is active as an NDPK 218 U/mg; as a negative control, we made a second protein nm23-H2-H118F mutation, killing the catalytic activity (Figure 2D).

Intact protein immobilized into magnetic beads was directly analyzed by mass spectrometry to obtain information on the exact molecular mass. The protein was subjected to MALDI mass spectrometry, producing the spectrum shown in Figure 2E. A single sharp peak was detected in the mass spectrum, exhibiting an average molecular mass of 20920.3 Da. An aliquot of nm23-H2 was then digested with trypsin, and the resulting peptide mixture was directly analyzed by MALDI mass spectrometry, producing the spectrum shown in Figure 3C. The mass spectral analysis led to the verification of about 92% of the entire protein sequence (Figure 3B).

An additional assay was performed to determine if nm23-H2 in vitro can be phosphorylated by casein kinase I. A previous article (4) has already shown the potential of nm23 protein to be phosphorylated by casein kinase II.

In addition, using the predicting phosphorylated software available at (<http://www.cbs.dtu.dk/services/NetPhos/> and <http://www.cbs.dtu.dk/databases/PhosphoBase/>), we have observed a region around S120–S125 that could be a potential site for casein kinase I (Figure 3A).

The MALDI analyses showed the presence of three series of related peaks (Figure 3C). The other two series of doublets were assigned to the mono- and diphosphorylated forms of the protein isoforms. Serines 125, 122, and 120 were found phosphorylated as shown in trypsin digestion pattern (Figure 3C).

In this report, we show our capability to express nm23-H2 protein (20 kDa in size) in a 96-well format and purify it by using Ni-NTA agarose beads all the way by using a robotic workstation, developing an automated method coupled to mass spectrometry within a detailed quality-control analysis. To guarantee and increase the rate of purified protein, the method presented here uses chaperones recombinant (Hsp40 and Hsp70) baculovirus co-infection to ensure appropriate protein folding. The automation procedure runs in 3 h and

is a useful method for industrial and academic environments to produce large-scale phase-recombinant eukaryotic proteins to be assayed for function in a systematic manner. To date, our laboratory has used this method to purify several expressed proteins ranging from 20 to 70 kDa in size, with a recovery rate of 2–4 pmol purified protein/well, thus further confirming the reproducibility and consistency of the automated protocol.

In addition, this method will allow one to characterize the catalytic domain and/or the specific interaction sites in a protein complex; experiments described as of random and site directed mutagenesis, generating high number of mutants, would be expressed at the same time and tested for their enzymatic activity. Finally, several tests have been performed to verify the quality of the recombinant protein produced, including NDPK activity assays and in vitro phosphorylation experiments to further validate the quality and usefulness of the method developed.

ACKNOWLEDGMENTS

We thank Monika Destruelle (Qia-gen) and Guido Meneghini (Perkin Elmer-Packard BioScience, Italy) for technical support. This work was supported by an AIRC-FIRC Research Fellowship (A.D.), a 2002 AIRC-FIRC grant (M.Z.), a Compagnia San Paolo Torino 2002 grant (M.Z.), a 2001 TIGEM-Telethon Regione Campania grant (M.Z.), and a PRIN 2002 grant MIUR-Rome (G.M. and M.Z.).

REFERENCES

1. Kost, T.A. and J.P. Condreay. 1999. Recombinant baculoviruses as expression vectors for insect and mammalian cells. *Curr. Opin. Biotechnol.* 10:428-433.
2. Hartl, F.U. and M. Hayer-Hartl. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295:1852-1858.
3. Yokoyama, N., M. Hirata, K. Ohtsuka, Y. Nishiyama, K. Fujii, M. Fujita, K. Kuzushima, T. Kiyono, and T. Tsurumi. 2000. Co-expression of human chaperone Hsp70 and Hsp40 co-factor increases solubility of overexpressed target proteins in insect cells. *Biochim. Biophys. Acta* 1493:119-124.

4. Engel, M., O.G. Issinger, I. Lascu, T. Seib, S. Dooley, K.D. Zang, and C. Welter. 1994. Phosphorylation of nm23/nucleoside diphosphate kinase by casein kinase 2 in vitro. *Biochem. Biophys. Res. Commun.* 199:1041-1048.
5. Freije, J.M., P. Blay, N.J. MacDonald, R.E. Manrow, and P.S. Steeg. 1997. Site-directed mutation of Nm23-H1. Mutations lacking motility suppressive capacity upon transfection are deficient in histidine-dependent protein phosphotransferase pathways in vitro. *J. Biol. Chem.* 272:5525-5532.
6. Hartsough, M.T. and P.S. Steeg. 2000. Nm23/nucleoside diphosphate kinase in human cancers. *J. Bioenerg. Biomembr.* 32:301-308.
7. Postel, E.H., V.H. Weiss, J. Beneken, and A. Kirtane. 1996. Mutational analysis of NM23-H2/NDP kinase identifies the structural domains critical to recognition of a c-myc regulatory element. *Proc. Natl. Acad. Sci. USA* 93:6892-6897.
8. Lanio, T., A. Jeltsch, and A. Pingoud. 2000. Automated purification of His6-tagged proteins allows exhaustive screening of libraries generated by random mutagenesis. *BioTechniques* 29:338-342.
9. Dell, A. and H.R. Morris. 2001. Glycoprotein structure determination by mass spectrometry. *Science* 291:2351-2356.
10. Hoffmann, R., S. Metzger, B. Spengler, and L. Otvos, Jr. 1999. Sequencing of peptides phosphorylated on serines and threonines by post-source decay in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Mass Spectrom.* 34:1195-1204.
11. Kuster, B. and M. Mann. 1998. Identifying proteins and post-translational modifications by mass spectrometry. *Curr. Opin. Struct. Biol.* 8:393-400.
12. Zhou, H., J.D. Watts, and R. Aebersold. 2001. A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* 19:375-378.
13. Timmons, L., J. Xu, G. Hersperger, X.F. Deng, and A. Shearn. 1995. Point mutations in awdKpn which revert the prune/Killer of prune lethal interaction affect conserved residues that are involved in nucleoside diphosphate kinase substrate binding and catalysis. *J. Biol. Chem.* 270:23021-23030.
14. Hamby, C.V., R. Abbi, N. Prasad, C. Stauffer, J. Thomson, C.E. Mendola, V. Sidorov, and J.M. Backer. 2000. Expression of a catalytically inactive H118Y mutant of nm23-H2 suppresses the metastatic potential of line IV Cl 1 human melanoma cells. *Int. J. Cancer* 88:547-553.

Received 25 February 2003; accepted 27 March 2003.

Address correspondence to Massimo Zollo, Telethon Institute of Genetics and Medicine, Via Pietro Castellino 111, 80131 Naples, Italy. e-mail: zollo@tigem.it