SHIMADZU
Excellence in Science

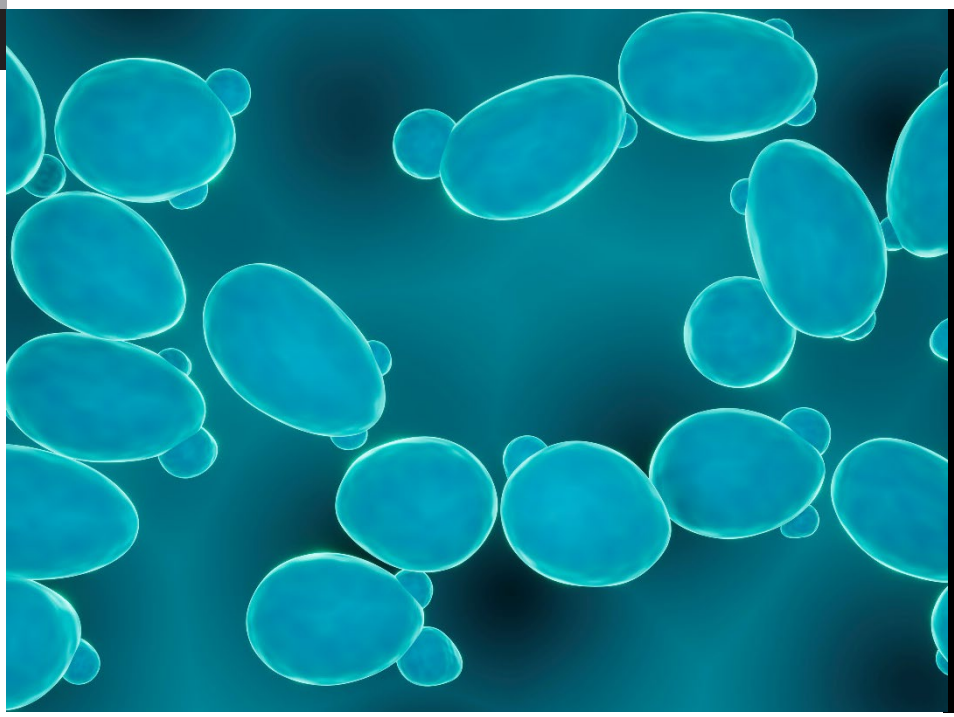
Application Note

No. 87

Life Sciences

Using Targeted Proteomics with an Ultra-Fast Triple Quadrupole Mass Spectrometer to Confirm Protein Overexpression

Fumio Matsuda[1,2,3] and Yutaka Umakoshi[4]



Life Science

■ Abstract

Protein overexpression is an experimental approach that is utilized in a variety of fields of research. Protein expression is most commonly confirmed with Western blotting, but this technique requires an antibody to detect each target protein. However, protein expression can be confirmed by targeted proteomics using a triple quadrupole mass spectrometer in MRM mode to selectively detect and quantify peptides produced by trypsin digestion of the target protein.

In this article, targeted proteomics was used to confirm protein expression by using an *E. coli* genetically modified to overexpress phosphoglucokinase (Pgk). The trypsin-digested samples used in the analysis were prepared by a proven experimental protocol. The targeted proteomics software Skyline was used to automatically create 110 MRM transitions that were based on the amino-acid sequence of Pgk, which it then exported to an LCMS-8060. With this data, the nano-LC-triple quadrupole mass spectrometer (LCMS-8060) was able to perform measurements that confirmed Pgk overexpression.

1. Introduction

Proteins undertake a wide range of functions in living organisms. Protein overexpression is an experimental strategy with many uses, including elucidating protein function in basic research, constructing modified microorganisms for bioproduction, and the biopharmaceutical production of antibodies and other proteins. Confirming whether overexpression has been achieved requires the measurement of protein expression levels, which are typically measured by Western blotting. However, this technique requires an antibody to detect each targeted protein. While antibodies for a variety of proteins are commercially available for use in basic medical research, in the field of metabolic engineering, genes from non-model animals are overexpressed in microorganisms and require the preparation of an antibody for each target protein.

1 Graduate School of Information Science and Technology, Osaka University
2 Osaka University Shimadzu Omics Innovation Research Laboratories
3 Industrial Biotechnology Initiative Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University
4 Solutions COE, Analytical & Measuring Instruments Division, Shimadzu Corporation

2. Confirming Pgk Overexpression in *E. coli* by Targeted Proteomics

To measure protein expression levels by targeted proteomics, a crude protein extract is digested with trypsin to create a mixture of trypsin-digested peptides. Next, the mixture of trypsin-digested peptides, which includes peptides from the target protein, is separated by reverse-phase liquid chromatography before the peptides are selectively detected and quantified using a triple quadrupole mass spectrometer in MRM mode. This method confirms protein overexpression without the need to prepare antibodies for the target protein. Trypsin digestion also generates a large number of peptides from the target protein. The amino-acid sequence of the target protein can then be used to configure specific MRM transitions for each peptide that predictions show will be generated by trypsin digestion of the target protein. As a very large number of MRM transitions are required, a triple quadrupole mass spectrometer that can operate at high speed should prove effective for this method.

Phosphoglucokinase (Pgk) is a key enzyme responsible for ATP generation in the glycolytic system. Currently, attempts are being made in the field of metabolic engineering to overexpress Pgk in *E. coli*. The study attempted to use targeted proteomics to confirm Pgk expression in *E. coli* that were engineered to overexpress Pgk.

E. coli K-12 and a Pgk-overexpressing *E. coli* strain from an *E. coli* overexpression library (ASKA library[1]) were cultured in LB media; expression was induced with isopropyl *β*-D-1-thiogalactopyranoside (IPTG); cells were recovered from the culture broth by centrifugation when the turbidity reached OD 1.0; and then proteins were extracted with a cell-solubilizing buffer. A method used to extract proteins from yeast was then applied to *E. coli* and other microorganisms with favorable results (Table 1).[2] Trypsin digestion was also performed according to a proven and published experimental protocol.[3] For cultured human cells, using a commercial kit is a convenient option for protein extraction and trypsin digestion (https://www.funakoshi.co.jp/contents/8038).

Table 1 Method for Extracting Proteins from Microorganisms[2,3]

1	Measure the absorbance of the liquid culture broth to confirm there are sufficient cells. (Determine the amount of cells needed to prepare a 300 µg/mL protein solution in advance.)
2	Transfer 15 mL of whole culture broth into a Falcon tube and centrifuge (3000 rpm for 5 min).
3	Dispose of the supernatant and add 1 mL of bacterial cell lysis buffer pre-cooled on ice (50 mM Hepes buffer [pH 7.5], 5 % glycerol, 15 mM dithiothreitol [DDT], 100 mM KCl, 5 mM EDTA, and 1 tablet of cOmplete Protease Inhibitor Cocktail).
4	Agitate the suspension thoroughly with a vortex mixer before transferring to a 1.5 mL tube and adding zirconia beads (6 mm × 1, 0.6 mm × enough to fill 1 PCR tube).
5	Homogenize for 2.5 min, and then cool on ice before homogenizing again for 2.5 min until the suspension becomes cloudy.
6	Centrifuge (15000 rpm for 5 min).
7	Transfer the supernatant to a 1.5 mL tube for proteomic analysis (low protein binding tube).
8	Measure the amount of protein using the Bio-RAD RC DC protein assay (Bio-Rad).

3. Automated MRM Method Creation Using Skyline Software

Targeted proteomics utilizes various rules of thumb. For example, trypsin-digested peptides of 6 to 16 amino acids are best suited to LC-MS analysis; electrospray ionization of trypsin-digested peptides generates divalent protonated molecules; and fragmentation by collision-induced dissociation mainly generates y-series fragment ions. These types of rules can be used along with the amino-acid sequence of the target protein to select trypsin-digested peptides best suited for LC-MS analysis and configure MRM transitions that selectively detect each of these peptides. However, doing this manually is very complicated.

Skyline is free software developed by the University of Washington for the Windows operating environment.[4] It implements almost a complete targeted proteomics workflow from method creation to data processing and can even directly read and write method and data files used by Shimadzu's triple quadrupole mass spectrometers.

The capability of the software to automatically generate MRM transitions from a protein amino-acid sequence was used in this research.[5] First, the amino-acid sequence of Pgk was obtained from a database in the FASTA format (Fig. 1a). The text data was copied and pasted into the "Target" window of Skyline. It then automatically selected 16 trypsin-digested peptides and then generated 110 MRM transitions to selectively detect those peptides (Fig. 1c). This data was then exported as an LC-MS method file (Fig. 1d). More information on this procedure can be found in Skyline tutorials.

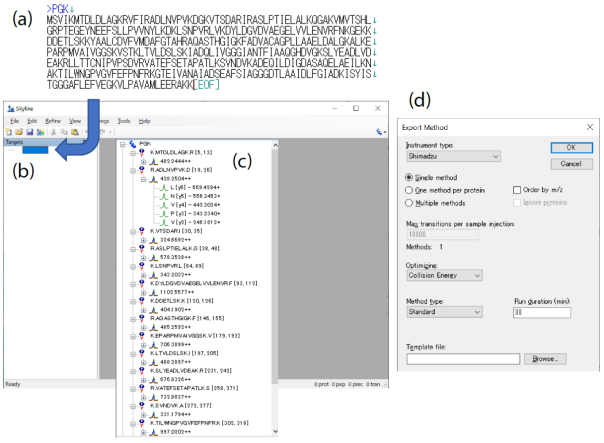


Fig. 1 Generating MRM Transitions Automatically with Skyline

4. Confirming Pgk Overexpression in *E. coli* Using a Nano-LC-Triple Quadrupole Mass Spectrometer

Trypsin-digested Pgk peptides prepared from *E. coli* K-12 and Pgk-overexpressing *E. coli* were analyzed using a nano-LC-triple quadrupole mass spectrometer (LCMS-8060). The equipment setup and the analytical conditions are shown in Table 2. Analyzing the Pgk-overexpressing *E. coli* sample produced multiple sharp and strong peaks for trypsin-digested Pgk peptides. The signal intensity of these peaks was substantially greater than the same peaks detected from the wild-type *E-coli* (K-12). Given the impossibility that a protein matching the MRM transitions for trypsin-digested Pgk peptides would be produced incidentally, this result confirmed that Pgk was overexpressed.

Furthermore, when there is sufficient leeway in terms of analytical sensitivity (Fig. 2) a semi-micro LC can be used for detection in place of a nano-LC, that is, targeted proteome analysis can be performed with a semi-micro LC triple quadrupole mass spectrometer, which is intended for low molecular weight compound analysis, without special adjustments to the equipment.

The results of Pgk overexpression analysis can also be used to select 3 or 4 trypsin-digested Pgk peptides that are detected with high sensitivity to create an MRM assay that measures Pgk expression. The group of authors that is responsible for this article is currently creating MRM assays for key *E. coli* and *Saccharomyces cerevisiae* proteins.[6,7]

Table 2 Nano-LC-Triple Quadrupole Mass Spectrometer Setup and Analytical Conditions

LC:	**LC-20ADnano**
Eluent A:	0.1 % formic acid + 5 % acetonitrile aq. solution
Eluent B:	0.1 % formic acid + 95 % acetonitrile aq. solution
Eluent C (for Trap):	0.1 % formic acid aq. Solution
Online Degasser:	One unit for eluent A and one for eluent B (to prevent mixing)
Piping:	nanoViper fingertight fitting (Thermo Scientific)
Flow Rate:	400 nL/min
Gradient:	B. conc. 0 % (0-7 min) – 65 % (45 min) – 100 % (50-65 min) – 75 % (67 min) - 0% (75-90 min) Injection valve switched to sample introduction side at 5 minutes.
Trap Flow Rate:	40 µL/min
Column Temperature:	Not controlled
Column:	L-column Micro L-C18 0.1 × 150 mm, 3 µm (CERI)
Trap Column:	L-column Micro L-C18 0.3 × 5 mm, 5 µm (CERI)
Tip:	Fortis tip 150-20 (AMR)
Interface:	nano-spray interface (AMR)
MS:	**LCMS-8060**
Nebulizer Gas Flow:	None
DL Temperature:	150 °C
Heat Block Temperature:	200 °C
Drying Gas Flow:	None
Q1 Resolution:	Low
Q3 Resolution:	Low
CID Gas:	270 kPa
Interface Voltage:	1.4 to 1.7 kV
Detection Method:	MRM mode

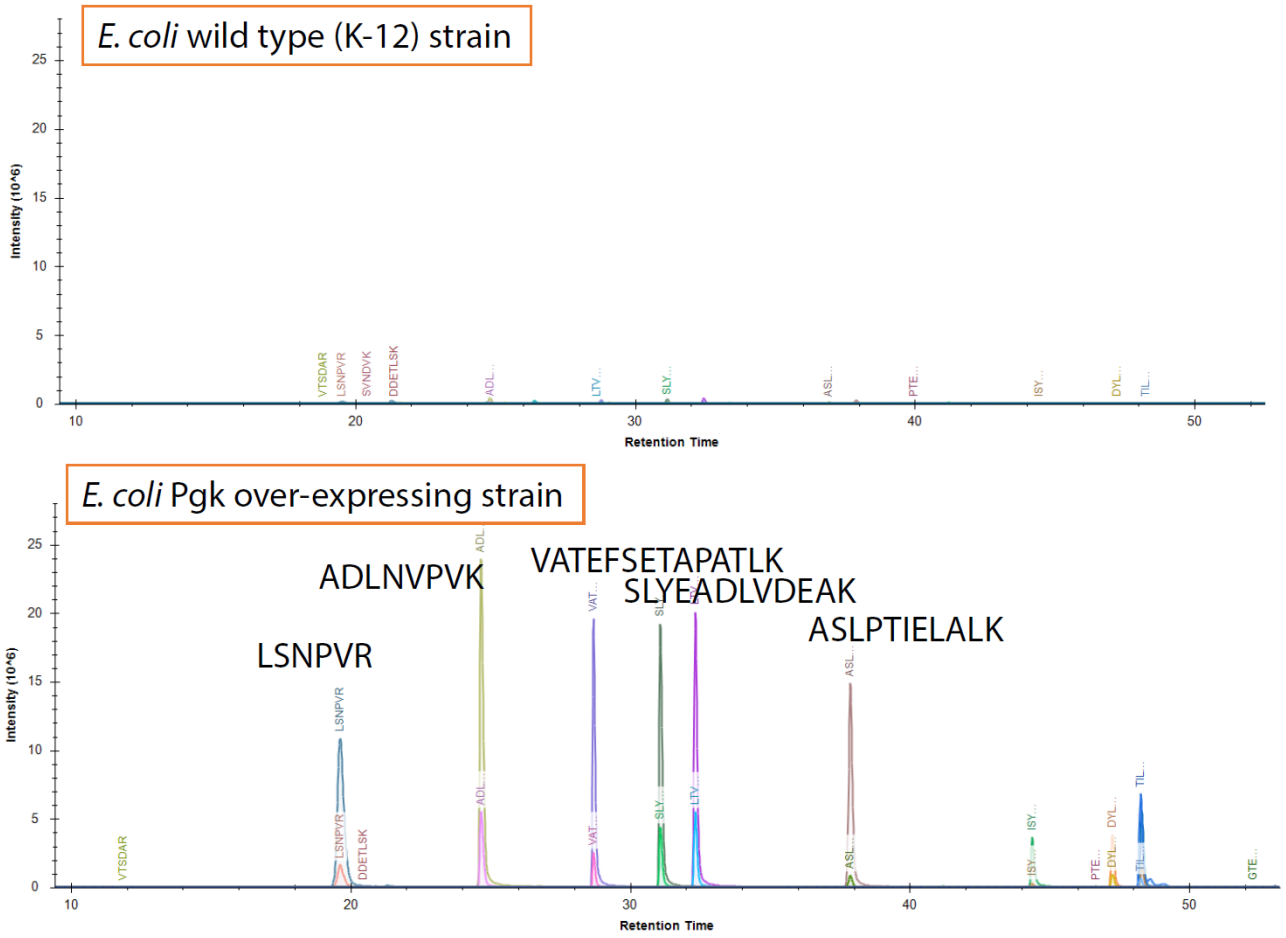


Fig. 2 Confirmed Pgk Expression by *E. coli*
16 trypsin-digested Pgk peptides measured by MRM method.

5. Conclusion

Metabolic engineering modifies microorganisms to overexpress multiple enzymes with the aim of creating microorganisms for use in bioproduction. Protein overexpression is also commonly used in basic medical research related to humans. Targeted proteomics, which does not require antibodies for protein detection, offers useful and rapid methods for confirming protein overexpression. Triple quadrupole mass spectrometers capable of high-speed analysis, can assist in the adoption of targeted proteomics, which requires large numbers of MRM transitions.

Acknowledgments

We would like to thank Etsuko Wakisaka, Hiroshi Kitano, and Atsumi Tomita of (at the time of this research) the Graduate School of Information Science and Technology, Osaka University, for their technical assistance with the proteome analysis performed in this study.

< References >

1) M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori. Complete set of ORF clones of Escherichia coli ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA research : an international journal for rapid publication of reports on genes and genomes* 12: 291-299, 2005.

2) Picotti P, Bodenmiller B, Mueller LN, *et al*. Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. Cell, 2009;138:795-806.

3) Uchida Y, Tachikawa M, Obuchi W, *et al*. A study protocol for quantitative targeted absolute proteomics (QTAP) by LC-MS/MS: application for inter-strain differences in protein expression levels of transporters, receptors, claudin-5, and marker proteins at the blood-brain barrier in ddY, FVB, and C57BL/6J mice. Fluids Barriers CNS, 2013; 10:21.

4) MacLean B, Tomazela DM, Shulman N, *et al*. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics, 2010; 26:966-968.

5) Skyline website: https://brendanx-uw1.gs.washington.edu/labkey/project/home/software/Skyline/begin.view (Accessed on October 10, 2023)

6) Matsuda, F., Tomita, A., Shimizu, H., 2017. Prediction of hopeless peptides unlikely to be selected for targeted proteome analysis. Mass Spectrometry. 6, A0056.

7) Uebayashi, K., Shimizu, H., Matsuda, F., 2018. Comparative analysis of fermentation and enzyme expression profiles among industrial *Saccharomyces cerevisiae* strains. Appl. Microbiol. Biotechnol. 102, 7071-7081.

First Edition: Dec. 2023

SHIMADZU
Shimadzu Corporation
www.shimadzu.com/an/