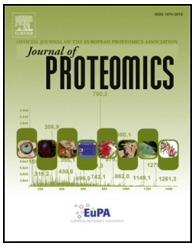
JOURNAL OF PROTEOMICS XX (2013) XXX–XXX

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot



Microwave-assisted acid hydrolysis of proteins combined with peptide fractionation and mass spectrometry analysis for characterizing protein terminal sequences☆

*Lu Chen, Nan Wang, Difei Sun, Liang Li**

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada

ARTICLE INFO

Keywords:
Protein sequencing
Terminal peptides
Microwave-assisted acid hydrolysis
Peptide fractionation
LC–MS/MS

ABSTRACT

We report a relatively simple mass spectrometric technique for characterizing the terminal amino acid sequences of proteins. It is based on the use of microwave-assisted acid hydrolysis (MAAH) with 3 M HCl to hydrolyze a protein into polypeptide ladders with varying sizes of up to the molecular mass of the protein. The hydrolysate is then fractionated by isocratic reversed phase liquid chromatography (RPLC) to produce a low-mass-peptide fraction mainly consisting of the terminal peptides. This fraction is subjected to LC tandem mass spectrometry (MS/MS) analysis to generate the terminal peptide sequence information. Using bovine serum albumin as an example, it is shown that more than 10 terminal peptides of each end could be identified using as little as 0.5 μg (7.5 pmol) of protein. This method was applied for the characterization of a recombinant protein (mCherry with an additional sequence tag added to the N-terminal for expression and purification) and its truncated form (mCherry treated with enterokinase to cleave off the tag). Sequence errors and unexpected by-products with different terminal sequences were determined from these two samples, illustrating that this method of HCl MAAH with peptide fractionation and LC–MS/MS analysis should be useful for detailed characterization of protein terminal sequences.

Biological significance
Protein terminal truncation or modification plays an important role in determining the biological functions of a protein. Detailed characterization of protein terminal sequences is critical in biological studies as well as in the development and quality control of protein-based therapeutics and vaccines. In this work, we report a relatively simple method for analyzing protein terminal sequences based on microwave-assisted acid hydrolysis to generate the peptide ladder of a protein, liquid chromatography fractionation of the resultant ladder to collect the low-mass-peptide fraction which mainly contains terminal peptides, and LC–ESI MS/MS sequencing of the collected peptides.
This article is part of a Special Issue entitled: Can Proteomics Fill the Gap Between Genomics and Phenotypes?

☆ This article is part of a Special Issue entitled: Can Proteomics Fill the Gap Between Genomics and Phenotypes?

* *Corresponding author at:* Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada. Tel.: +1 780 492 3250; fax: +1 780 492 8231.
E-mail address: Liang.Li@ualberta.ca (L. Li).

1874-3919/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jprot.2013.10.014

1. Introduction

Protein terminal truncation and other modifications are common post-translational modifications that can alter protein structures and functions [1–4]. In addition, protein production and storage of protein-based pharmaceutical preparations may also cause the terminal degradation, resulting in changes of drug efficacy [5]. Thus determining any changes of the protein termini is very important for biological studies as well as for quality control during the production of protein-based therapeutics and vaccines. Edman degradation has been traditionally used for protein N-terminal sequencing [6]. However, this method only applies to the proteins whose N termini are unblocked or unmodified at the N-terminal amino acid and the process is time-consuming. For protein C-terminal sequencing, Schlack–Kump degradation with (iso)thiocyanate [7,8] is a reaction similar to Edman degradation. The disadvantages of this method are low efficiency and low reproducibility and generally less than ten residues of protein C-terminus can be obtained. In comparison, MS-based methods are fast, sensitive and can identify protein modifications. In many areas of applications, MS has become a preferred choice for protein sequencing. To generate the terminal peptide sequence information, proteins can be digested by exopeptidases such as CPY and CPB [9,10] or chemicals such as cyanogens bromide (CNBr) [11]. The cleaved peptides can be analyzed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) or electrospray ionization (ESI) tandem mass spectrometry (MS/MS). However, protein digestion using an enzyme such as trypsin or a combination of multiple enzymes may fail to generate the terminal peptides with suitable sizes or at a sufficient abundance for successful MS sequencing [12]. In some cases, the quality of the MS/MS spectrum may not be sufficiently high for deducing information on individual amino acids or modified amino acids in the terminal peptides. Top-down proteomics provides another way for protein sequence analysis. MALDI in-source decay (ISD) [13], ESI electron capture dissociation (ECD) or electron transfer dissociation (ETD) [14,15] is proved to be efficient to identify terminal peptides of mid-size proteins. However, these methods usually cannot provide the individual amino acid sequence information on the terminal peptides, particularly for relatively large proteins. New strategies for selective capturing of the terminal peptides from a proteomic digest that are subsequently sequenced by LC–MS/MS have been developed for relatively high throughput analysis of protein termini [2,16–19].

One alternative approach that is particularly suitable for detailed characterization of a protein sequence is based on the use of microwave-assisted acid hydrolysis (MAAH) to degrade a protein into many short peptides, followed by MS and MS/MS analysis of the resultant peptides [20–23]. The sequences of these short peptide ladders are often overlapped to some extent, thereby a complete picture of the amino acid linkage or any modifications therein can be deduced [24]. This shotgun protein sequencing method can cover all or a large portion of individual amino acids in a sequence. However, in some cases, the terminal peptide sequences may be missing due to relatively low abundance of these peptides generated, compared to the internal peptides, and ion suppression during the detection. For internal sequence, if a small set of peptide ladders covering a stretch of amino acids are not detected by MS due to their low abundances or ion suppression, other sets of more detectable overlapping peptide ladders can cover these amino acids. However, for terminal peptides, if a set of terminal peptides are not detected by MS, no other more detectable peptide ladders can cover this region.

In this work, we describe an improved shotgun sequencing method for characterizing the terminal peptides. To increase the likelihood of detecting the terminal peptides, we developed a rapid HPLC fractionation method to fractionate the hydrolysate of a protein generated by MAAH using 3 M HCl. The low molecular weight peptides collected were found to contain many terminal peptides that could be sequenced using LC–MS/MS. This method was developed using bovine serum albumen (BSA) as a model protein. Both MALDI-TOF and LC–MS/MS techniques were examined for detecting the terminal peptides. This method was then applied to the analysis of a recombinant protein (a red fluorescent protein mCherry with a short sequence tag added to the N-termini) and its N-terminal truncated form (mCherry minus the tag) to illustrate how this method can be used to characterize a protein or protein product in a real world situation.

2. Material and methods

2.1. Chemicals and reagents

Dithiothreitol (DTT), iodoacetamide (IAA), ammonium bicarbonate (NH_4HCO_3), trifluoroacetic acid (TFA), LC–MS grade formic acid (FA), bovine serum albumin were purchased from Sigma-Aldrich Canada (Markham, ON, Canada). LC–MS grade water and acetonitrile (ACN) were from Fisher Scientific Canada (Edmonton, AB, Canada). ACS grade 37% HCl was from Merck (KGaA, Darmstadt, Germany).

2.2. mCherry protein preparation

Electrocompetent *E. coli* strain DH10B (Invitrogen) was transformed with plasmid containing the gene of the target protein (pBAD-mCherry) and plated on Lysogeny Broth (LB) medium (1% tryptone (BD 211705), 0.5% yeast extract (BD 212750), and 1% NaCl (Caledon 7560-1 from Sigma-Aldrich) agar plates supplemented with 0.1 mg/ml ampicillin (Fisher BP1760-25) and 0.02% L-arabinose (Sigma A3256). Then the plates were incubated for overnight at 37 °C.

The cell culture and preparation of the cell lysates of *E. coli* were similar to those reported [25] with some modifications. Five mL LB culture of mCherry cells from the LB was incubated at 37 °C with shaking at 225 rpm overnight. The culture was centrifuged at 3901 *g* for 10 min. The pellets were resuspended in 50 mL of LB and incubated at 37 °C with shaking at 225 rpm overnight. Then the culture was centrifuged at 3901 *g* for 10 min. The pellets were resuspended in 500 mL of LB. Cells were incubated at 37 °C with shaking at 225 rpm for overnight. After cooling the medium to 4 °C, cells were centrifuged at 11,300 g for 20 min at 4 °C. The pellets were resuspended in 10 mL lysis buffer (50 mM sodium phosphate of pH 8.0, 300 mM NaCl 10 mM imidazole, 0.05 mg/mL DNase I (Roche 10 104 159

Please cite this article as: Chen L, et al, Microwave-assisted acid hydrolysis of proteins combined with peptide fractionation and mass spectrometry analysis for characterizi..., J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.10.014