Published in final edited form as:

Electrophoresis. 2015 August ; 36(15): 1690–1698. doi:10.1002/elps.201400579.

Improved protease digestion conditions for membrane protein detection

Lie Min, Leila H. Choe, and Kelvin H. Lee*

Department of Chemical and Biomolecular Engineering and Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711, USA

Abstract

This work presents improved protease digestion conditions for membrane protein detection. The enzymatic digest of bacteriorhodopsin (BR), a model membrane protein with seven transmembrane domains (TMDs) was investigated. An initial in-gel digestion identified 17% BR sequence coverage, including part of the seventh TMD. To improve sequence coverage, BR digest was tested with different concentrations of RapiGest, methanol (MeOH) and SDS using either trypsin or chymotrypsin. Two improved conditions, 0.01% SDS or the combination of 10% MeOH and 0.01% RapiGest, were chosen. Trypsin digestions in both conditions achieved more than 40% BR sequence coverage compared to 0% using standard digestion conditions. Peptides detected from trypsin and chymotrypsin digestions in the same condition were combined to maximize sequence coverage. The same conditions were applied to a different membrane protein with one TMD, Selenoprotein S, and proteins from *Escherichia coli*. For Selenoprotein S, a higher sequence coverage, including a peptide from the TMD, was detected from the improved condition compared to the typical condition. The application of both improved conditions to a membrane protein fraction of *Escherichia coli* resulted in the identification of 309 (SDS) and 329 (MeOH/RapiGest) unique proteins of which 140/309 and 148/329 were membrane proteins.

Keywords

Bacteriorhodopsin; *Escherichia coli*; membrane protein; proteomics

1 Introduction

Membrane proteins play important roles in many biological processes, including cell signal transduction, cell-cell communication, molecular transport, and cell motility [1]. The functions of membrane proteins are essential for maintaining normal biological processes and malfunction of membrane proteins often leads to various diseases [2]. It has been estimated that membrane proteins constitute 50–60% of drug targets on the market [2]. However, studies on membrane proteins are difficult because of low abundance and poor solubility. For example, membrane proteins are predicted to make up about 25% of all annotated proteins [3], but represent less than 2% of protein entries in the protein data bank

Correspondence: Dr. Kelvin H. Lee, Department of Chemical and Biomolecular Engineering and Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711, USA, KHL@udel.edu, Fax: +1 302 831 4841.

The authors have declared no conflict of interest.

(PDB) database [4]. In addition, they are typically under-reported in many large-scale proteomic publications [5].

As a highly sensitive tool with wide dynamic range, MS has been widely used for qualitative and quantitative characterization of proteins [6]. But in MS-based proteomic studies, most digestion conditions are optimized for soluble proteins. Membrane proteins are difficult to extract, digest, and detect [5, 7]; therefore, it is important to develop simple but robust methods for MS-based membrane protein analysis.

Different strategies have been developed to improve membrane protein solubility, including the addition of detergents or organic solvents [5, 7, 8]. Among commonly used detergents, SDS is one of the most powerful solubilization reagents. The concentration of SDS used in shotgun proteomics varies from 0.1% to up to 1%. Nonetheless, SDS negatively affects protease activity, interferes with LC separation and has poor MS compatibility. Thus, SDS is typically removed or diluted prior to digestion and MS/MS data acquisition; however, SDS has high affinity for proteins and peptides and is very difficult to remove. Alternatively, MS compatible, acid labile detergents, such as RapiGest (Waters), PPS Silent Surfactant (Protein Discovery), and the noncleavable detergent Invitrosol (Invitrogen) have been developed and are commercially available. Chen *et al.* compared the effectiveness of these three detergents and found the highest number of unique proteins and peptides were obtained with the aid of RapiGest [9, 10]. Upon acidification, RapiGest decomposes into a MS compatible product and a water immiscible product; the latter can be removed by centrifugation. However, at high detergent concentrations, hydrophobic peptides can co-precipitate with the water immiscible product, which can lead to severe sample loss. An additional potential drawback of the acid labile detergents is a relative high cost, which can limit extensive application for membrane proteomic analysis.

Organic solvents have also been reported to assist in the solubilization of membrane proteins [10, 11]. Compared to the above-discussed detergents, organic solvents can be easily removed after protease digestion. Zhang *et al.* and Blonder *et al.* used 60% MeOH to improve the solubility of the *Escherichia coli* membrane proteome and human epidermal membrane proteome [11–13]. Nevertheless, due to the limited solubilization ability of organic solvents, large amounts of starting materials are needed. Furthermore, organic solvents at high concentrations could dramatically inhibit protease activity.

In this work, protease digestion conditions using low concentration and different combinations of additives that can assist membrane protein solubility, while interfering minimally with protease digestion and MS, were developed. The respective effects of MeOH, RapiGest and SDS on trypsin and chymotrypsin activity were tested. The digestion of BR was tested in a range of concentrations of MeOH, RapiGest and SDS with either trypsin or chymotrypsin. The sequence coverages from trypsin or chymotrypsin digestions in different conditions were compared. To maximize sequence coverage, peptides detected from trypsin or chymotrypsin digestion were combined. The improved conditions were applied to a second model membrane protein with one TMD – Selenoprotein S and to a well-characterized organism, *Escherichia coli* to demonstrate the general applicability.

2 Materials and methods

2.1 Preparation of *Halobacterium salinarum* purple membrane

The purple membrane was prepared from *Halobacterium salinarum* according to literature [14, 15]. Briefly, *H. salinarum* were grown in 1 L of rich medium composed of 250 g NaCl (Fisher Scientific, Fair Lawn, NJ), 20 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ (Fisher Scientific), 3 g $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7 \cdot 5\text{H}_2\text{O}$ (Fisher Scientific), 2 g KCl (Fisher Scientific), and 10 g peptone oxid (Sigma-Aldrich, St. Louis, MO) at 39 °C with constant agitation and illumination in an Erlenmeyer flask. When OD_{600} reached 0.6, agitation was stopped and the cells were grown for 3 more days. Then the cells were harvested by centrifugation at 3000 g for 15 min at 4 °C, lysed with osmotic lysis in deionized water, and incubated at room temperature for 30 min in the presence of a small amount of deoxyribonuclease (Sigma-Aldrich). The cell lysate was centrifuged at 3000 g at 4 °C for 10 min to remove cell debris. Then the supernatant was subjected to ultracentrifugation (Optima™ Ultracentrifuge, Beckman Coulter, Brea, CA) at 30000 g at 4 °C for 30 min to enrich membrane proteins. The pellets were washed with 0.1 M NaCl 3 times; then the resultant pellet was resuspended in deionized water and stored at −80 °C until further use. Such preparations of purple membrane contain mostly BR [16]. The concentration of BR was quantified with the following formula: $[c] = \text{OD}_{568\text{nm}} / 62700 \times 26000 \text{ (mg/mL)}$ [17, 18].

2.2 Protease activity determination

Activities of trypsin (Promega, Madison, WI) or chymotrypsin (Roche, Indianapolis, IN) were measured in triplicate using a Fluorescent Detection Kit (Sigma-Aldrich) per manufacturer's instructions. Protease activities in the presence of different concentrations of MeOH (Mallinckrodt Chemicals, Phillipsburg, NJ), RapiGest (Waters, Milford, MA) and SDS (Bio-Rad Laboratories, Hercules, CA) were normalized to protease activities without any additives. Fluorescence was measured on a SpectraMax M5 microplate reader (Molecular Devices, Sunnyvale, CA).

2.3 In-gel digestion and analysis of BR

Samples containing 5 µg BR were loaded to a precast 12% Tris-HCl Ready Gel (Bio-Rad Laboratories), and run at 150 V constant voltage in a Mini Protean 3 Cell (Bio-Rad Laboratories) until the dye front reached the bottom of the gel. The gel was stained with SYPRO Ruby Protein Gel Stain (Molecular Probes, Eugene, OR, USA), and destained per manufacturer's instruction. In-gel digestion of BR was performed similarly as previously reported [19]. Briefly, 1.5mm gel pieces were excised and digested in a Progest (Digilab, Marlborough, MA). Samples were subjected to reduction/alkylation with DTT (Bio-Rad Laboratories) and iodoacetamide (Sigma-Aldrich). Enzymatic digestion was performed with trypsin (Promega) at 37 °C. Subsequently, samples were desalted and concentrated using Ziptips (Millipore) and applied to target plates with alpha-cyano-4-hydroxycinnamic acid matrix (Sigma-Aldrich). Data were collected on a 4800 MALDI TOF-TOF Analyzer (AB Sciex, Foster City, CA) in positive ion, reflector mode over a mass range of 850–4000 m/z, with internal calibration. Select peaks were further analyzed by MSMS at 1 kV with default calibration. Combined MS and MSMS data were submitted to Mascot v2.2 (Matrix Science, London, UK) and searched against NCBI, with trypsin specificity, 100ppm mass tolerance,

0.3Da MSMS tolerance, and the following variable modifications: carbamidomethylation (C) and oxidation (M).

2.4 In-solution digestion of BR and Selenoprotein S

For standard conditions, enzymatic digestion of BR and Selenoprotein S (10 µg) was carried out with either trypsin or chymotrypsin in 50 mM ammonium bicarbonate (ABC, Fisher Scientific) buffer without any sonication. In the improved method, the same amount of model protein was sonicated on ice in 50 mM ABC containing 0%, 10% or 30% MeOH, without detergent, or with 0.01% RapiGest, 0.01% SDS, or 0.1% SDS. The sonication was performed with a cup horn sonicator (Misonix, Farmingdale, NY) set to 42 W output energy. Samples were subsequently subjected to reduction with 100 mM DTT for 25 min at 95 °C followed by alkylation with 150 mM iodoacetamide for 30 min at room temperature in the dark. Although BR does not contain disulfide bonds, the reduction step was incorporated to keep the protocol consistent. After a second sonication on ice for 5 min, enzymatic digestions were carried out with trypsin or chymotrypsin at 37 °C for 16 hours at an enzyme/substrate mass ratio of 1:20. The digest was then acidified with 10% formic acid (Acros Organics, Fair Lawn, NJ), and incubated at room temperature for 2 hours followed by centrifugation at 16000 g for 5 min.

2.5 Preparation and digestion of *Escherichia coli* proteins

Luria-Bertani (LB, Fisher Scientific) broth (3 mL) was inoculated with *E. coli* W3110 cells; cultures were grown overnight with shaking at 37 °C. One mL of overnight culture was used to inoculate 50 mL fresh LB in an Erlenmeyer flask. After OD₆₀₀ reached 0.6, the cells were harvested by centrifugation at 3000 g for 10 min at 4 °C. Cell pellets were washed 2 times with an ice-cold solution containing 10 mM Tris (Fisher Scientific), 75 mM NaCl and 1 mM EDTA (Fisher Scientific), then sonicated in the same buffer on ice for 5 min. The cell lysate was centrifuged at 2500 g for 5 min. The resultant supernatant was further centrifuged at 173000 g on an ultracentrifuge. The resultant pellet was used as membrane protein fraction.

The *E. coli* membrane protein fractions were resuspended in 50 mM ABC buffer without detergent or with 0.01% SDS or with 10% MeOH + 0.01% RapiGest. Protein suspensions were agitated at 37 °C for 30 min, sonicated for 5 min on ice, and assayed for protein concentration with a Bradford assay.

Samples with 50 µg *E. coli* membrane protein fractions were digested the same way as the model proteins.

2.6 Sample Cleanup – Detergent Removal

After digestion of model proteins (BR and Selenoprotein S), the DetergentOUT™ GBS10-125 detergent removal kit (G-Biosciences, St. Louis, MO) was used to remove excess detergents per manufacturer's instructions.

2.7 RP-LC and LC-MS/MS analysis

LC-MS/MS analysis of BR and Selenoprotein S was performed on a QTrap 2000 (AB Sciex) equipped with an Agilent 1100 HPLC (Agilent, Wilmington, DE). Digests were

injected onto a C18 reverse phase column (3 μm , 100 $\text{\AA}$, 100 mm $\times$ 2 mm, Phenomenex, Torrance, CA). Gradient elution was performed with mobile phase A (0.1% formic acid in water) and B (0.1% formic acid in ACN (Mallinckrodt Chemicals)) at a flow rate of 200 $\mu\text{L}/\text{min}$. A program of 2% B for 10 min, 2% to 50% B in 26 min, 50% to 98% B in 1 min, and 98% B for 3 min was used to elute peptides. The eluted peptides were ionized with a Turbo V source and data were acquired at 5500 V and 450 $^{\circ}\text{C}$, with 400–1500 m/z survey scan set to trigger enhanced resolution scans and up to 3 enhanced product ion scans with dynamic exclusion, and a scan range of 100–1500 m/z with rolling collision energy in positive ion mode.

Digests of *E. coli* membrane protein fractions were fractionated on an Agilent 1100 HPLC equipped with a RP-LC-S reverse phase column (3 μm , 100 $\text{\AA}$, 150 mm $\times$ 2.1 mm, Varian Inc., Palo Alto, CA). Gradient elution was performed with the mobile phase C (50 mM NH_4OH (J.T. Baker Company, Phillipsburg, NJ) in 2% ACN) and mobile phase D (50 mM NH_4OH in 80% ACN). A program of 4% D for 10 min, 4% to 10% D in 2 min, 10% to 50% D in 28 min, 50% to 96% D in 10 min, and 96% D for 10 min was used to elute the peptides. The elution was monitored with UV absorbance at 215 nm. Fifteen fractions were collected and subsequently dried by vacuum concentration. Each fraction was reconstituted in 32 μL of 2% ACN in 0.1% formic acid and 15 μL was analyzed on a QTrap 4000 (AB Sciex) equipped with an UltiMate 3000 HPLC system (Dionex, Sunnyvale, CA). Samples were injected onto a C18 precolumn (PepMap100, 300 μM I.D. $\times$ 5 mm, 5 μM , 100 $\text{\AA}$, Dionex). Peptides were washed with mobile phase E (0.1% formic acid in 2% ACN) for 5 min, then eluted to a C18 column (Acclaim PepMap100, 75 μm I.D. $\times$ 150 mm, 3 μm , 100 $\text{\AA}$, Dionex) through a program of 2% mobile phase F (0.1% formic acid in 98% ACN) for 5 min, 2% to 10% F in 10 min, 10% to 50% F in 65 min, 50% to 90% F in 10 min, 90% F for 10 min at a flow rate of 250 nL/min. The eluted peptides were ionized with a Nanospray II source (AB Sciex). Data were acquired at 2500 V, 150 $^{\circ}\text{C}$ with 450–1800 m/z MS survey scans set to trigger enhanced resolution scans and up to 6 enhanced product ion scans with dynamic exclusion and scan range of 100–1800 m/z in positive ion mode.

2.8 Data analysis for in-solution digestion and shotgun proteomics

For model proteins, the MS/MS data were submitted to Mascot.dll v1.6 from Analyst (v1.6.1, AB Sciex) and searched against NCBI nr with the following parameters: *Archaea* taxonomy (for BR) or *Homo sapiens* (for Selenoprotein S); trypsin or chymotrypsin as specified enzymes allowing for up to 2 missed cleavages; carbamidomethylation (Cys) and oxidation (Met) as variable modifications; peptide tolerance of 1.2 Da, MS/MS tolerance of 0.6 Da. For BR, pyro-Glu (N-term Glu and Gln) was specified as an additional variable modification. Protein identifications were based on matches with at least 95% confidence interval. Average sequence coverage of BR was calculated from triplicate experiments.

E. coli membrane protein MS/MS data were submitted for searches, against a local copy of an *E. coli* database downloaded from NCBI, through ProteinPilot software (v3, AB Sciex). One search was performed using Mascot v2.2 search engine. Proteins identified from the Mascot search were used for the majority of the analysis, including categorizing membrane proteins and subcellular location, identifying transmembrane domains, and calculating grand

average of hydropathy (GRAVY) [20] scores and protein sequence coverage was obtained from ProteinPilot's Paragon algorithm. Integral membrane proteins were predicted with the transmembrane hidden Markov model algorithm [21] (<http://www.cbs.dtu.dk/services/TMHMM/>). Gene names for protein hits were matched from protein accession numbers using iProClass (<http://pir.georgetown.edu/pirwww/dbinfo/iproclass.shtml>). Subcellular locations of *E. coli* proteins were assigned based on gene annotations in EchoBASE, a database for *E. coli* [22]. For the few proteins without matched gene names in EchoBASE, subcellular locations were predicted with the online server PSORTb (v3.0.2, <http://www.psort.org/psortb/>). The GRAVY scores of proteins were calculated in batch with the following online server: <http://www.gravy-calculator.de/>.

The *E. coli* membrane protein MS/MS data have been deposited to the ProteomeXchange Consortium [23] via the PRIDE partner repository with the dataset identifier PXD001824.

3 Results and Discussion

3.1 Effects of additives on protease activities

To establish a baseline of digestion in standard conditions, BR was in-gel digested with trypsin followed by MALDI-TOF/TOF detection. In total, 17% sequence coverage was achieved, including part of the 7th TMD. Sequence coverage and the number of detected TMDs were also evaluated by in-solution digestion of BR with trypsin and chymotrypsin.

The effects of multiple solubility enhancers on the protease activities of trypsin and chymotrypsin were tested at various concentrations. It was found that high levels of MeOH, SDS and RapiGest had negative impacts on trypsin and chymotrypsin activities (Figure 1). However, 10% MeOH did not have a negative effect on enzyme activities for either enzyme, while 30% MeOH resulted in a 30% reduction in activities of both enzymes. 0.01% SDS had a moderate negative impact on enzyme activity (40% reduction for both enzymes) compared to 0.1% SDS (90% reduction). 0.01% RapiGest had no significant negative impact on trypsin activity and resulted in a 30% reduction for chymotrypsin activity. 0.1% RapiGest yielded a 30% reduction in enzyme activity for trypsin and a 40% reduction for chymotrypsin. The results were consistent with the findings of Masuda *et al.* in which 10% MeOH enhanced trypsin activity, but 80% MeOH or 0.1% SDS or 0.1% RapiGest inhibited activity [24].

3.2 Optimization of digestion conditions

Based on this initial screening, trypsin and chymotrypsin digestions were tested with a range of enhancer concentrations (0%, 10%, 30% MeOH with either 0% or 0.01% RapiGest) using BR as the substrate. With the standard digestion conditions using either trypsin or chymotrypsin, no BR peptides with 95% confidence or above were detected. When digested with MeOH and/or RapiGest and sonication, high confidence BR peptides were detected. In the enhanced trypsin digestions, average BR sequence coverages from triplicate experiments were similar for all 6 conditions (around 15%) (Figure 2A). In the enhanced chymotrypsin digestions, BR sequence coverage reached a maximum (24%) in 10% MeOH + 0.01% RapiGest (Figure 2A). To compare the overall effectiveness of each condition, peptides detected from the same method with trypsin and chymotrypsin were pooled and reported.

The optimum condition was 10% MeOH + 0.01% RapiGest (subsequently referred to as “MeOH/RapiGest”), with a combined BR sequence coverage of 52% (Figure 2B).

In addition, digestion was tested in 0.01% and 0.1% SDS with either trypsin or chymotrypsin using BR as the substrate. In 0.1% SDS, no high confidence BR peptides were detected. In 0.01% SDS, a range of sequence coverage, from 0% to 28%, was achieved (Figure 2C).

Although SDS is one of the most powerful solubilization solvents, it is well known to have detrimental effects on RP-LC and MS, possibly obscuring MS results [25]. SDS has been reported to broaden peaks widths, affect peak shape and alter elution time in RP-LC [26]. In MS, SDS could form peptide-SDS adduct, mask peptide signal, and suppress peptide ionization [27]. To further improve upon the best-case SDS digestion, a sample clean-up step was incorporated to minimize SDS interference. Average BR sequence coverage improved from 12% to 46% in the 0.01% SDS (subsequently referred to as “SDS”) trypsin digestion with sample clean-up (Figure 2C). Average sequence coverage from chymotrypsin digestion did not significantly improve with sample clean-up. Sample clean-up was also applied to BR digestions in MeOH/RapiGest. The results are similar to those from the digestions in SDS. With sample clean-up, average sequence coverage from trypsin digestion improved from 20% to 41%, and average sequence coverage from chymotrypsin did not significantly improve (Figure 2D).

To study the overall enhancement, peptides detected from the samples having undergone clean-up in the same method with trypsin and chymotrypsin are combined for reporting. In SDS with sample clean-up, BR sequence coverage with trypsin and chymotrypsin reached 66%; in MeOH/RapiGest with sample clean-up, it reached 65% (Figure 2E), compared to 0% from the standard digestion conditions. The results here indicate that the two improved digestion conditions worked similarly at solubilizing membrane proteins.

To further reveal the differences of the two improved digestion conditions, the detected BR TMDs were examined and are listed in Table 1. Without sample clean-up and in 0.01% SDS, no TMD peptides were detected from trypsin digestion and parts of the second and third TMDs were detected from chymotrypsin digestion; whereas in MeOH/RapiGest, the first TMD was detected with trypsin, and parts of the second and third, the fourth and sixth TMDs were detected with chymotrypsin. With sample clean-up, the first, fifth and the seventh TMDs were detected in both enhanced conditions with trypsin; and parts of the second and third, and the sixth TMDs were detected from chymotrypsin.

Membrane proteins, especially integral membrane proteins with TMDs, are very hydrophobic and therefore challenging for typical proteomics workflows. On the one hand, membrane proteins need the addition of solubilization reagents to get into solution; on the other hand, these additives typically negatively impact protease digestion and MS detection. BR has 7 TMDs and is often used as a model membrane protein because of its hydrophobic nature. With the two improved conditions and sample clean-up, BR sequence coverages from trypsin alone were significantly improved (more than 40% vs 0%) compared to the standard digestion conditions. Although trypsin has been the protease of choice in MS-based

proteomics [28], membrane proteins, especially TMDs, often contain fewer trypsin cleavage sites, and trypsin digestion of membrane proteins often leads to low protein sequence coverage [1, 29]. However, this study shows that by combining results from trypsin and chymotrypsin digestion, using the improved conditions and incorporating sample clean-up, sequence coverage of BR could be maximized, and peptides from each of the 7 TMDs could be detected.

3.3 Application of the improved digestion conditions

3.3.1 Application of the improved digestion conditions to a different

membrane protein with one TMD—To test the applicability of the two enhanced conditions, SDS and MeOH/RapiGest, digestion of Selenoprotein S with trypsin and chymotrypsin was performed and compared with standard digestion conditions. Sample clean-up was applied. Under standard digestion conditions, Selenoprotein S sequence coverages from trypsin and chymotrypsin reached 46% and 25% respectively, with combined sequence coverage reaching 54%; in SDS and MeOH/RapiGest, combined sequence coverages were 54% and 63% respectively (Figure 3). In MeOH/RapiGest, not only were more sequences detected but also a peptide from the TMD was detected whereas under standard conditions no TMD peptides were detected. This indicates that the two conditions could not only be applied to the digestion of very hydrophobic membrane protein such as BR (with GRAVY score of 0.725), but also less hydrophobic membrane protein such as Selenoprotein S (with GRAVY score of -0.629). Improved digestion conditions may be most beneficial for more hydrophobic proteins as seen in the cases of BR (Figure 2E) vs. Selenoprotein S (Figure 3).

3.3.2 Application of the improved digestion conditions to *E. coli* lysate

—To test the applicability of the improved digestion conditions across a broader range of proteins, an *E. coli* membrane protein fraction was digested in the two enhanced conditions, SDS and MeOH/RapiGest. For comparison, the membrane protein fraction was also digested with the standard digestion conditions. The unique protein hits, membrane protein hits, and integral membrane protein hits with 1–18 TMDs were compared (Figure 4A and 4B). For the trypsin digestions of the membrane protein fraction, the number of unique protein hits (Figure 4A), membrane protein hits (Figure 4B), and integral membrane protein hits (Figure 4B, the fraction under divider) all increased from the standard digestion conditions to the two detergent-based conditions. This trend is also seen for the chymotrypsin digestion (Figure 4A and Figure 4B).

GRAVY scores can be used to evaluate protein hydrophobicity [20]. Proteins were classified into four groups according to their GRAVY scores: <-0.5 , $-0.5-0$, $0-0.5$, >0.5 (Figure 4C); proteins with GRAVY scores between $0-0.5$ are hydrophobic and >0.5 are very hydrophobic; proteins with GRAVY scores between $-0.5-0$ are hydrophilic and <-0.5 are very hydrophilic. It can be seen that for both trypsin and chymotrypsin digestions, a predominant proportion of proteins identified have calculated GRAVY scores between $-0.5-0$. With trypsin, the percentages of the identified hydrophobic proteins and very hydrophobic proteins in the two detergent-based solutions were similar (16% and 11% proteins with GRAVY scores between $0-0.5$ and >0.5 , respectively). With chymotrypsin,

about 15% proteins with GRAVY scores >0.5 were detected in MeOH/RapiGest. Wu *et al.* studied the *E. coli* membrane proteome with 1% RapiGest and 1% SDS and identified about 12% proteins with GRAVY scores >0.5 [30]. The percentage of very hydrophobic proteins obtained by Wu *et al.* was similar to what was achieved here. However, a much lower concentration of SDS (0.01%) and a combination of 0.01% RapiGest and 10% MeOH were used in this study. The concentrations of additives used in our work are significantly lower than most commonly reported additive concentrations [9–13, 24, 30, 31].

To further compare the improved conditions to the standard digestion condition, sequence coverages from the two improved conditions were normalized to the sequence coverages from the standard digestion condition and plotted against protein GRAVY scores. Only proteins co-identified from the standard digestion condition and either one of the improved conditions are shown in Figure 4D and 4E. Normalized sequence coverage at one indicates that similar sequence coverages were found from the improved conditions and from the standard digestion condition. Normalized sequence coverage >1 indicates that more peptide sequences were identified from the improved conditions. As shown in Figure 4D (trypsin digestion), in the MeOH/RapiGest method, two of the proteins with GRAVY scores >0.5 show normalized sequence coverages of about one, and the rest of the proteins all show normalized sequence coverages above one. In Figure 4E (chymotrypsin digestion), one of the proteins with GRAVY score >0.5 showed normalized sequence coverage of about one, and the rest proteins all show values above one in the SDS method. The data here support that the two improved conditions help sequence coverages for hydrophobic proteins. The GRAVY scores against sequence coverages for all identified proteins were plotted in Supplementary Figure 1.

To compare the membrane protein profile observed in this study with a reference proteome, the proteins detected from the two improved conditions with trypsin and chymotrypsin were sorted by subcellular location, and the distribution of detected proteins across subcellular locations was compared to the distribution of subcellular locations of all proteins in *EchoBASE* (Figure 4F). *EchoBASE* is an integrative *E. coli* post-genomic database and contains an annotation of the *E. coli* genome [22]. Based on subcellular location, identified proteins were classified into the following groups: cytoplasmic, inner membrane, outer membrane, periplasmic, extracellular, and unknown. The percentages of proteins at each subcellular location were compared. The subcellular distributions of membrane fraction proteins from the two improved conditions with either trypsin or chymotrypsin digestion were very similar to each other. The percentages of the cytoplasmic proteins from the improved conditions were significantly lower than that in *EchoBASE* (49%, 49% vs 66%), while the percentages of the outer membrane proteins were much higher than those in the *EchoBASE* (20%, 20% vs 3%). The trend was the same for chymotrypsin digestion. These observations are consistent with improved solubilization and digestion of membrane proteins in the improved conditions.

Profiles of the number of TMDs per protein for the integral membrane proteins identified were compared to the integral membrane proteins in *EchoBASE* (Figure 4G). The overall profiles of TMDs per protein from the two improved conditions (with either trypsin or chymotrypsin) were similar to each other and were similar to the distribution profile in

EchoBASE. Integral membrane proteins with more than 14 TMDs were identified with trypsin digestion in both enhanced conditions and the percentages of these integral membrane proteins were higher than that in *EchoBASE*. It is noteworthy that in the literature, integral membrane proteins identifications often contain an unusually high percentage of proteins with 1 TMD and a much lower percentage of proteins with more than 1 TMD [30–33], while the distribution by number of TMDs of integral membrane proteins in the current study was similar to the distribution annotated in *EchoBASE*. This observation indicates that the two improved digestion conditions provided increased detection of membrane proteins.

4. Concluding remarks

Two improved protease digestion conditions using low concentration and different combination of additives were identified for membrane protein detection. The improved conditions achieved more than 40% BR sequence coverage, as compared to 0% in standard digestion conditions. When applied to *E. coli* membrane protein fractions, both conditions provided similar numbers of unique protein identifications for total and for membrane proteins. The two conditions are relatively easy to implement and can be applied to high-throughput proteomic studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by NIH 5P30RR031160-02. We thank Dr. Tzvetana Lazarova at the Universitat Autònoma de Barcelona (Spain) for providing *H. salinarum*. We thank Dr. Sharon Rozovsky at the University of Delaware for providing Selenoprotein S.

Abbreviations

ABC	ammonium bicarbonate (NH ₄ HCO ₃)
MeOH	methanol
BR	bacteriorhodopsin
TMDs	transmembrane domains

References

1. Macher BA, Yen TY. Mol Biosyst. 2007; 3:705–713. [PubMed: 17882332]
2. Hopkins AL, Groom CR. Nat Rev Drug Discov. 2002; 1:727–730. [PubMed: 12209152]
3. Fagerberg L, Jonasson K, von Heijne G, Uhlen M, Berglund L. Proteomics. 2010; 10:1141–1149. [PubMed: 20175080]
4. Fau KE, Punta M, Rost B. Curr Opin Struct Biol. 2012; 22:326–332. [PubMed: 22622032]
5. Vuckovic D, Dagley LF, Purcell AW, Emili A. Proteomics. 2013; 13:404–423. [PubMed: 23125154]
6. Domon B, Aebersold R. Science. 2006; 312:212–217. [PubMed: 16614208]
7. Helbig AO, Heck AJR, Slijper M. J Proteomics. 2010; 73:868–878. [PubMed: 20096812]

8. Chen EI, McClatchy D, Park SK, Yates JR III. *Anal Chem.* 2008; 80:8694–8701. [PubMed: 18937422]
9. Chen EI, Cociorva D, Norris JL, Yates JR III. *J Proteome Res.* 2007; 6:2529–2538. [PubMed: 17530876]
10. Zhang N, Chen R, Young N, Wishart D, Winter P, Weiner JH, Li L. *Proteomics.* 2007; 7:484–493. [PubMed: 17309111]
11. Blonder J, Conrads TP, Yu L, Terunuma A, Janini GM, Issaq HJ, Vogel JC, Veenstra TD. *Proteomics.* 2004; 4:31–45. [PubMed: 14730670]
12. Mitra SK, Gant JA, Ruby JF, Clouse SD, Goshe MB. *J Proteome Res.* 2007; 6:1933–1950. [PubMed: 17432890]
13. Stoeckenius DOW. *Methods Enzymol.* 1974; 31:667–678. [PubMed: 4418026]
14. Becher BM, Cassim JY. *Prep Biochem.* 1975; 5:161–178. [PubMed: 1144313]
15. Rietschel B, Bornemann S, Arrey TN, Baeumlisberger D, Karas M, Meyer B. *Proteomics.* 2009; 9:5553–5557. [PubMed: 20017156]
16. Turner GJ, Chittiboyina S, Pohren L, Hines KG, Correia JJ, Mitchell DC. *Biochemistry.* 2009; 48:1112–1122. [PubMed: 19140737]
17. Fau RM, Heyn MP. *Biochemistry.* 1979; 18:4977–4983. [PubMed: 508727]
18. Choe LH, Aggarwal K, Franck Z, Lee KH. *Electrophoresis.* 2005; 26:2437–2449. [PubMed: 15924362]
19. Kyte J, Doolittle RF. *J Mol Biol.* 1982; 157:105–132. [PubMed: 7108955]
20. Krogh A, Larsson B, Heijne GV, Sonnhammer ELL. *J Mol Biol.* 2001; 305:567–580. [PubMed: 11152613]
21. Misra RV, Horler RSP, Reindl W, Goryanin II, Thomas GH. *Nucleic Acids Res.* 2005; 33:D329–D333. [PubMed: 15608209]
22. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, Dienes JA, Sun Z, Farrah T, Bandeira N, Binz P-A, Xenarios I, Eisenacher M, Mayer G, Gatto L, Campos A, Chalkley RJ, Kraus H-J, Albar JP, Martinez-Bartolome S, Apweiler R, Omenn GS, Martens L, Jones AR, Hermjakob H. *Nat Biotech.* 2014; 32:223–226.
23. Masuda T, Tomita M, Ishihama Y. *J Proteome Res.* 2008; 7:731–740. [PubMed: 18183947]
24. Rundlett KL, Armstrong DW. *Anal Chem.* 1996; 68:3493–3497. [PubMed: 21619282]
25. Swiderek KM, Alpert AJ, Heckendorf A, Nugent K, Patterson SD. *ABRF: News Method and Reviews.* 1997 Dec.:17–25.
26. Loo RR, D N, Andrews PC. *Protein Sci.* 1994; 3:1975–1983. [PubMed: 7703844]
27. Vandermarliere E, Mueller M, Martens L. *Mass Spectrom Rev.* 2013; 32:453–465. [PubMed: 23775586]
28. Eichacker LA, Granvogl B, Mirus O, Müller BC, Miess C, Schleiff E. *J Biol Chem.* 2004; 279:50915–50922. [PubMed: 15452135]
29. Wu F, Sun D, Wang N, Gong Y, Li L. *Anal Chim Acta.* 2011; 698:36–43. [PubMed: 21645657]
30. Lin Y, Lin H, Liu Z, Wang K, Yan Y. *J Sep Sci.* 2014; 37:3321–3329. [PubMed: 25196059]
31. Lin Y, Huo L, Liu Z, Li J, Liu Y, He Q, Wang X, Liang S. *PLoS One.* 2013; 8:e59779. [PubMed: 23555778]
32. Lin Y, Liu H, Liu Z, Wang X, Liang S. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2012; 901:18–24.

Figure 1. Effects of additives on protease activity. (1A) Effects of MeOH; (1B) Effects of SDS and RapiGest.

Figure 2.

Optimization of BR digestion conditions. (2A) Sequence coverages from either trypsin or chymotrypsin digestion, and (2B) combined sequence coverages from single protease digestion results in the following conditions: (1) 0% MeOH, (2) 10% MeOH, (3) 30% MeOH, (4) 0% MeOH + 0.01% RapiGest, (5) 10% MeOH + 0.01% RapiGest, (6) 30% MeOH + 0.01% RapiGest. (2C) Sequence coverages with either trypsin or chymotrypsin in 0.01% SDS (SDS) with or without sample clean-up; (2D) Sequence coverages with either trypsin or chymotrypsin in 10% MeOH + 0.01% RapiGest (MeOH/RapiGest) with or without sample clean-up; (2E) Combined sequence coverages from single protease digestions in different digestion conditions with sample clean-up.

Figure 3.
Combined sequence coverages of Selenoprotein S from single protease digestions in different conditions.

Figure 4.

E. coli proteomic study. (4A, 4B) Unique proteins and membrane proteins identified from the membrane protein fraction (MPF) under standard digestion condition, SDS, and MeOH/RapiGest with trypsin and chymotrypsin digestions; in 4B, fractions above the dividers and below the dividers represent the number of membrane proteins without and with TMDs, respectively; (4C) distributions of the GRAVY scores among identified proteins; (4D, 4E) scatter plot of normalized protein sequence coverages against GRAVY scores with trypsin or chymotrypsin; (4F) distribution of subcellular localizations of proteins from the two

enhanced conditions with trypsin and chymotrypsin digestions compared to those in *EchoBASE*; (4G) TMDs per protein profiles from the two enhanced conditions compared to those in *EchoBASE*.

Table 1

Bacteriorhodopsin TMDs detected in different digestion conditions

Additives	Sample clean-up	Protease used	Detected TMD	Peptide sequence	GRAVY score
0.01% SDS	No	Trypsin	None	N/A	N/A
0.01% SDS	No	Chymotrypsin	Part of 2nd Part of 3rd	YAITTLVPAIAF GEQNPIYWARY	1.742 -1.291
0.01% SDS	Yes	Trypsin	1st 5th Part of 7th	XAQITGRPEWIWALGTALMGLGTLVFLVK VYSYRFVWWAISTAAML YIL YVLFFGFTSK VGFLILLR	0.743 1.063 1.956
0.01% SDS	Yes	Chymotrypsin	Part of 1st Part of 2nd Part of 3rd 6th	FLVKGMGVSDPDAKKF AITTLVPAIAF GEQNPIYWARY KVLNRNVTVVLWSAYPVVW	-0.025 2.018 -1.291 0.917
MeOH/RapiGest	No	Trypsin	1st Part of 2nd	XAQITGRPEWIWALGTALMGLGTLVFLVK YAITTLVPAIAF	0.743 1.742
MeOH/RapiGest	No	Chymotrypsin	Part of 3rd 4th 6th	GEQNPIYWARY AIVGADGLMIGTGLVGALTKAY KVLNRNVTVVLWSAYPVVW	-1.291 1.173 0.917
MeOH/RapiGest	Yes	Trypsin	1st 5th Part of 7th	XAQITGRPEWIWALGTALMGLGTLVFLVK VYSYRFVWWAISTAAML YIL YVLFFGFTSK VGFLILLR	0.743 1.063 1.956
MeOH/RapiGest	Yes	Chymotrypsin	Part of 2nd Part of 3rd 6th	YAITTLVPAIAF GEQNPIYWARY VLRNVTVVLWSAYPVVW	1.742 -1.291 1.2