

Acoustic technology-assisted rapid proteolysis for high-throughput proteome analysis

Bora Kim¹, Trang Tran Huyen¹, Na-Young Han¹, Jong-Moon Park¹,
Ungsik Yu¹ and Hookeun Lee^{1,2,*}

¹Lee Gil Ya Cancer and Diabetes Institute, Gachon University, Songdo-dong,
Yeosu-ku, Incheon 406-840, Korea

²College of Pharmacy, Gachon University, Songdo-dong, Yeosu-ku, Incheon 406-840, Korea

(Received September 23, 2011; Revised November 18, 2011; Accepted November 18, 2011)

대량 발굴 프로테오믹 분석을 위한 어쿠스틱 기술 기반 고속 단백질 절편화

김보라¹ · 천후엔창¹ · 한나영¹ · 박종문¹ · 유웅식¹ · 이후근^{1,2,*}

¹가천대학교 이길여암당뇨연구원, ²가천대학교 약학대학
(2011. 9. 23. 접수, 2011. 11. 18. 수정, 2011. 11. 18. 승인)

Abstract: Recent developments and improvements of multiple technological elements including mass spectrometry (MS) instrument, multi-dimensional chromatographic separation, and software tools processing MS data resulted in benefits of large scale proteomics analysis. However, its throughput is limited by the speed and reproducibility of the protein digestion process. In this study, we demonstrated a new method for rapid proteolytic digestion of proteins using acoustic technology. Tryptic digests of BSA prepared at various conditions by super acoustic for optimization time and intensity were analyzed by LC-MS/MS showed higher sequence coverage in compared with traditional 16 hrs digestion method. The method was applied successfully for complex proteins of a breast cancer cells at 30 min of digestion at intensity 2. This new application reduces time-consuming of sample preparation with better efficiency, even with large amount of proteins, and increases high-throughput process in sample preparation state.

요약: 최근 질량분석기와 다차원 크로마토그래피 분석법과 대용량 데이터를 처리하는 생물정보학 프로그램 등과 같은 복합적인 기술요소들의 발전은 대량 프로테오믹 분석하는 것을 가능하게 했다. 하지만, 이런 대량의 프로테오믹 분석하는 것은 단백질 절편화 과정의 긴 소요 시간과 낮은 재현성으로 인하여 한계를 가진다. 이 연구에서는 어쿠스틱 기술을 이용해 빠르게 단백질을 분해하는 새로운 방법을 제시했다. 이 어쿠스틱 기술의 시간과 강도를 최적화하기 위해서 여러 가지 조건에서 BSA 단백질을 트립신으로 절편화한 후 액체 크로마토그래피와 질량분석기로 분석하였다. 16시간동안 인큐베이션하는 기존의 방법과 슈퍼 어쿠스틱 기술을 사용한 방법을 비교하였을 때 슈퍼 어쿠스틱 기술이 기존의 방법보다 더 높은 아미노산 동정율을 보였으며, 유방암 세포와 같은 단백질 복합체에 적용하였을 때 30분의 절편화 시

★ Corresponding author

Phone : +82-(0)32-899-6584 Fax : +82-(0)32-899-6519

E-mail : hklee@gachon.ac.kr

간에서도 효율적인 결과를 확인할 수 있었다. 이 새로운 방법은 샘플 전처리 과정에서 많은 양의 단백질 일지라도 더 좋은 효율성을 가지게 되고 또한, 소비되는 시간도 줄여주며 일정 시간내의 샘플 처리량도 증가하게 된다.

Key words : LC/MS/MS, protein digestion, acoustic technology

1. Introduction

Since the early 1990s, the term of “Proteomics” was coined by Marc Wilkins. Proteomics attempts to explain the information contained in gene sequences by systematic analysis of the proteins expressed in a cell or tissue. A prevailing method of proteomics analysis is as following: Proteins mixture which extracted from the raw samples is fragmented by proteolytic enzymes, the resulting peptides mixture is separated and subsequently analyzed by mass spectrometry (MS).^{1,2}

During the last decade, the robust development of analytical proteomics techniques including MS, multi-dimensional chromatographic separation, and bioinformatics software for MS data processing has benefited faster and higher-throughput protein identification and characterization.³⁻⁵ However, a significant bottleneck in the analysis pipeline is the sample-preparation stage, which is very important but has no dominant breakthroughs for a long time. Until now, sample preparation is still the most time-consuming and non-reproductive procedure in the analysis workflow.

Following the principle of MS instrument, protein sequencing by using collisionally induced dissociation is best accomplished by sequencing peptides instead of intact proteins.⁶ Therefore, the first step in the process of sample preparation is to divide the protein of interest into a representative series of peptides. For a long time, enzymatic hydrolysis is the most common method for this purpose.⁷ Each protease has its own unique action mechanism, in which, trypsin, specifically hydrolyzes peptide bonds at the carboxyl side of Lysine (K) and Arginine (R) residues, except when those bonds are to Proline (P) residues. Interestingly, the spacing of K and R residues in many proteins is such that many of the resulting

peptides are of a length well-suited to MS analysis. In addition, the protonation of those two sites which are two ends of the peptide makes high charged state, which is advantageous in MS ionization process. As a consequence, trypsin is by far the most widely used protease in proteomic studies.⁷

The traditional digestion protocol can take up to several hours to overnight because it requires extended incubation times.⁸⁻⁹ It is a result of ensuring completed hydrolysis when using low concentration enzyme to avoid the auto-digestion of trypsin, which might produce undesired trypsin fragments and complicate the protein complex.¹⁰

To improve the problem of long-time-consuming digestion processes, in recent years, several approaches have been developed for rapid-proteolytics. A promising approach was studied so far is microwave-assisted protein enzymatic digestion (MAPED).¹¹ Another strategy that aims at off-line proteolysis uses laser radiation which requires specific MALDI-TOF MS instrument.¹²

Immobilized trypsin is one of the earliest methods was applied in order to reduce auto-digestion of trypsin, allowing higher concentration in using trypsin and so require shorter incubation time. Moreover, it is very easy to remove the immobilized enzyme from peptides mixture prior to MS analysis, eliminating the “noise” of MS results.⁹ On the other hand, trypsin which is fixed on the static surface results in the easier interactions between enzyme and target proteins, increase the catalytic function. There are many reports indicated the possibility of protein digestion using trypsin immobilized on various support, including porous silicon matrix,¹³ glass,¹⁴ gel bead,¹⁵ polymer,¹⁶ sol-gel supports,¹⁷ membrane,¹⁸ and porous monolithic materials.¹⁹ These methods process online protein digestion during liquid chromatography with a

high pressure proteolytic reactor which denatured, digested and separated product peptides in a manner of minutes. Although it has many advantages and incubation time for complete digestion reduces as well, the issues of this approach is associated with clogging solid trypsin bound matrix or enzyme stability.²⁰ Moreover, small sample-loading capacity is the drawback of every online system in detection low-abundance target proteins. As a consequence, it limits the large-scale clinical proteomics, especially post-translational modification studies.

In this study, we would like to introduce a new method for rapid proteolytic digestion of protein using acoustic technology called The Covaris Adaptive Focuses Acoustics (AFA). This system was developed by Covaris technological company (USA). Similar to sonication, Covaris AFA also extracts proteins from cell base on acoustic process. Several keys make Covaris AFA to become a stronger tool in comparison with sonication are more controllable wavelengths and isothermal ability. Therefore, at the same time, it can not only create samples homogenization very quickly but also protect proteins mixture from degradation.

The most prominent characteristic of this system is sending acoustic energy wave packets from a dish-shaped transducer that converges and focuses to a small-localized area. At this focal point, the energy density may be controllably focused into the sample of interest which has proven to be beneficial to numerous applications of sample preparation. Sample vessel is placed in the water bath. The water bath temperature is controlled that the circulator will pump a temperature-controlled fluid through a stainless steel loop immersed in the water bath. Essentially, the process enables mechanical energy to be applied to a sample without directly contacting the sample.²¹

In associated with in-solution proteolysis in different time conditions, we demonstrate that acoustic technology offers rapid digestion in just only 30 minutes with higher sequences coverage. This new application reduces time-consuming of sample preparation, even with large amount of proteins, and increases high-throughput process in sample preparation state.

2. Experimental

2.1. Equipment set-up

The water bath holds approximately 1.6 liters of distilled or deionized water. The temperature range is from 5 °C to 40 °C which is controlled by hooking up an external recirculating chiller to the ports in the back of the system. A temperature sensor is to ensure the isothermal of water. Sample is added into sample vessel, which is placed in the water bath via sample holder. Acoustic transducer induces acoustic energy wave for protein extraction and rapid digestion. A computer which is set up with a specific software is connected directly to the super acoustic system. The main panel configures up to four different treatments as user wants (*Fig. 1*).

2.2. Materials and reagents

Sequencing-grade trypsin was purchased from Promega (Madison, WI). Six standard proteins (Bovine serum albumin (BSA), α -amylase, Transferrin, Myoglobin, β -lactoglobulin, β -casein), iodoacetamide (IAA), ammonium bicarbonate, trizma base and formic acid (FA) were obtained from Sigma-Aldrich (St. Louis, MO). Urea was ordered from Amresco (Solon, OH). TCEP and BCA assay kit were purchased from Pierce (Rockford, IL). HPLC grade solvents were bought from J. T. Baker (Phillipsburg, NJ). C18 ultra micro spin columns were purchased from Harvard

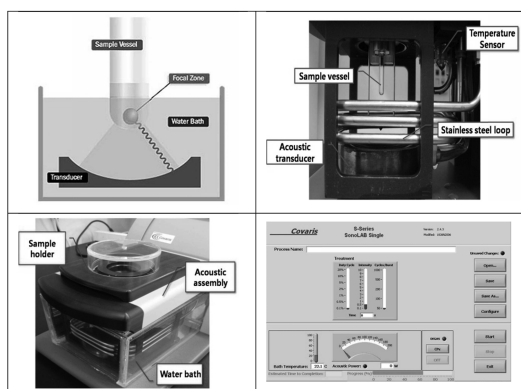


Fig. 1. The super acoustic system operation. The components include a water bath with temperature controller function, sample vessel and connected computer.

apparatus (Holliston, MA). Breast cancer cell line HS578T was quantitated by BCA method and analyzed as application data.

2.3. Digestion procedures

Protein samples were first resuspended by 8 M Urea and reduced with 5 mM TCEP at 37 °C during 30 min and alkylated at room temperature for 1 hour in the dark with 15 mM IAA. Sample solution could be split into several units with the same volume for an experiment in different conditions. Tris pH 8.3 was used to dilute concentration of Urea to 2 M and adjust pH of solution from 8 to 9 for trypsin reaction. Digestion was performed by adding trypsin into the protein solution at a substrate-to-enzyme ratio of 25:1 (w/w) either by current incubation method at 37 °C overnight or Covaris instrument. In the case of Covaris instrument, digestion was performed at different times (5, 10, 30, 45, and 60 min) or different intensity, at 37 °C, duty cycle 10%, cycle/burst 200, and 60 sec/cycle. The pH of the sample solutions was immediately adjusted between 2 and 3 after digestion step by adding formic acid to stop the trypsin activity. The peptide samples were cleaned up using C18 commercial micro spin column following company's protocol

2.4. LC-MS/MS analysis

Dried peptides were resuspended by 0.1% FA and separated in a capillary column packed with C18 at a flow rate of 0.3 μ L/min. we used an LC-MS/MS system with reverse phase (RP)-LC comprised of a Surveyor MS pump (Thermo Electron, San Jose, CA) for LC, a Spark auto sampler (EMMEN, The Netherlands), and a Finnigan LTQ linear ion trap mass spectrometer (MS) (Thermo Electron, San Jose, CA) equipped with nanospray ion (NSI) sources. The gradient was started at 2% B and a linear gradient to 60% B was achieved in 53 min, and then ramped to 80% B in 7 min and to 100% A over the next 30 min. The mass spectrometer was operated in a data-dependent mode (in a range of m/z 400-1600) in which each full MS scan was followed by five MS/MS scans where the five most abundant peptide

molecular ions were dynamically selected from the prior scan for collision-induced dissociation (CID) using a normalized collision energy of 35%. The temperature of the heated capillary and electrospray voltage were at 200 °C and 2.1 kV, respectively. To verify the reproducibility of our acoustic method, peptide samples were analyzed by 6520 Accurate-Mass Q-TOF LC/MS (Agilent Technologies, DE) with the same LC gradient at an electrospray potential of 1.8 kV. A drying gas flow of 5 L/min, and the fragmentor of 175 V were used. The qTOF was set to perform data acquisition in the positive mode that the m/z range of 100-3000 was used in the MS/MS scan.

2.5. Data analysis

The raw MS/MS data obtained from the LC-MS/MS analysis were format-converted by Trans-Proteomics Pipeline. Protein identification was performed using Computational Proteomics Analysis System (CPAS) search engine, version 9.1, and Thermo Proteome Discoverer, version 1.1.0.263. Trypsin was selected as protein cleavage specificity. Database searching of BSA was performed in International Protein Index (IPI) 17 standard mixtures (ipi.protein17mix.fasta). Six standard proteins mixture was searched using The Standard Protein Mix Database from Institute for Systems Biology (<http://regis-web.systemsbio.net/PublicDatasets/>). The complex protein samples were run against the IPI human database (version 3.52). Database search parameters included dynamic modification for oxidation of Met and for carbamidomethylation of Cys. The database search was performed with precursor and peptide mass tolerance of 50 ppm and 20 ppm respectively. PeptideProphet and Protein Prophet were used to evaluate the quality of peptide and protein identification with error rate less than 1%.

3. Results and Discussion

3.1. Optimization

Rapid proteolysis by the acoustic system was compared to the traditional incubation method by

LC-MS/MS analysis for BSA. In *Fig. 2*, the base peak chromatograms of BSA digests were shown for the conditions of different digestion time. Although there are some differences in peak patterns between normal and rapid digestion methods, the similar profiles of chromatogram including same main peaks

are achieved, even at 5 minutes digestion in the super acoustic system. It indicates that highly common peptides were formed despite the approximately 200 fold difference in processing time and the rapid digestion by the Covaris device was effectively carried out.

However, there are distinctive differences in the

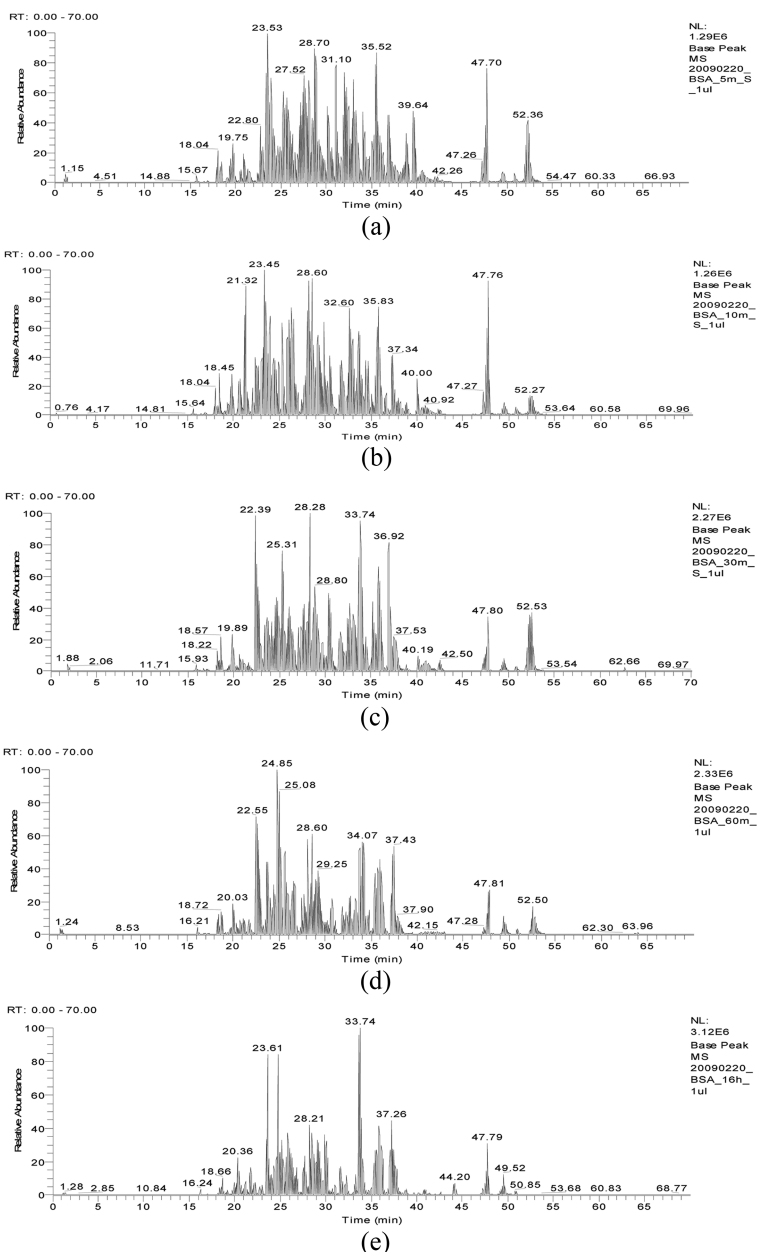


Fig. 2. Base peak chromatograms of BSA digests obtained at different digestion time-conditions. Samples were digested with Covaris instrument at the same intensity (0.5) in the period of (a) 5 min, (b) 10 min, (c) 30 min, and (d) 60 min, and (e) 16 hr incubation at 37 °C respectively.

peak areas of chromatograms, increased peak areas as digestion time decreases. This indicates that the digestion efficiency by the acoustic device is not as high as the normal digestion process, and generates significant amounts of peptides having missed cleavages. To systematically investigate this problem, MS/MS spectra were extracted from the chromatograms

and analyzed by X!Tandem search engine, and the results are shown in *Fig. 3*. The numbers of identified peptides at different time conditions were similar, even the rapid methods showed better results at 30 minutes or 60 minutes (*Fig. 3a*). Additionally, digested peptides provided from Covaris system in most cases (10 minutes, 30 minutes, 60 minutes) show higher percent of BSA sequence coverage in comparison with 16 hours digestion. Especially, at 30 minutes condition, the histogram represents the maximum coverage, 85.7%, much higher than in traditional incubation ($p < 0.05$) (*Fig. 3b*). It suggested that the super acoustic technology can be a useful protocol especially in post-translational modification study requiring high percent in coverage sequence.

However, one drawback of rapid digestion as predicted in *Fig. 2* and shown in *Fig. 3* is the high rate of missed cleavages generation, about 50% instead of 20% in the normal method. This finding assumes that the complete digestion perhaps cannot be achieved by Covaris at intensity 0.5. Though it has no effects in peptide identification, quantitative proteomics measuring relative peptide abundance changes should not be processed in this condition, because proteins quantification basically requires highly reproducible peptide generation for accurate quantitative measurements.

In order to find out the solution for missed cleavages generation problem and optimize conditions for rapid digestion protocol using Covaris system, proteolytic efficiency was tested by changing acoustic intensity in the range of 0.5 to 4. The same sets of BSA were digested by 16 hours incubation or new method which was set at different intensity conditions: 0.5, 1, 2, 3, 4 and 6 for 30 minutes.

As shown in *Fig. 4a*, among ranges of intensities parameter, the number of identified peptides results at intensity 2 or 3 is better with similar number of unique peptides but higher ratio of unique peptides over total identified peptides (60/885 and 53/522 respectively). Especially, these ratios are much higher than in conventional incubation method. In addition, digestion efficiency was still guaranteed by approximate percent of sequence coverage at different conditions

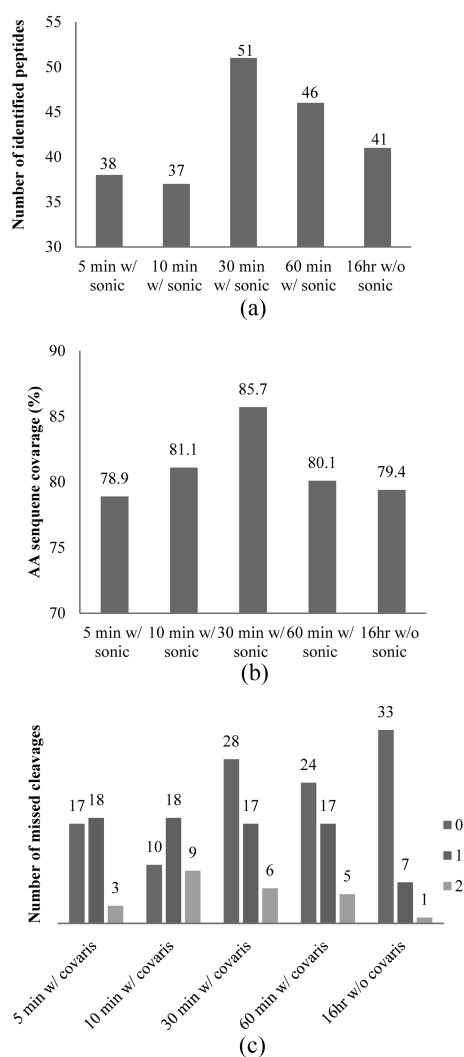


Fig. 3. Comparison of digestion efficiency of traditional and acoustic approach. Samples were digested with Covaris instrument at the same intensity (0.5) in the period of 5 min, 10 min, 30 min, and 60 min, and 16 hr incubation at 37 °C respectively. (a) Number of identified peptides, (b) BSA sequence coverage (%), (c) Number of missed cleavages, using X! Tandem search engine ($p < 0.05$).

(Fig. 4b). Considering missed cleavages problems, it is definitely improved by increasing the intensity from 0.5 to 2-4, and the consequence is the similar results between new and the normal digestion approach, suggesting for complete fragmentation (Fig. 4c).

On the other hand, it assumes that the number of unique peptides decreases when the intensity increases. It is due to missed cleavages were perfectly cut at cleavage sites into smaller fragments. The cleavage sites that remain at low intensity were cut off and new peptides have same sequenced with identified peptides, results in decreasing the number of unique peptides.

In overall, the condition of intensity 2 or 3 for 30 minutes is demonstrated as optimized conditions for Covaris system setting in rapid acoustic digestion which

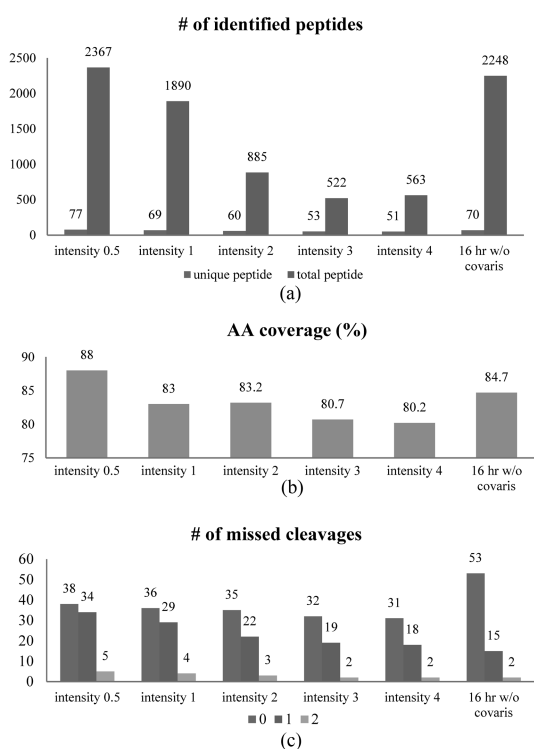


Fig. 4. BSA digestion by Covaris system at: Duty cycle: 10%, Time: 30 min, Cycle/burst: 200, intensity conditions: 0.5, 1, 2, 3, and 4 respectively, or at 16 hours incubation. MS/MS data were searched by X!Tandem program. (a) The number of identified peptides, (b) Percent of amino acid sequence coverage, (c) The number of missed cleavages at different digestion condition.

is applied effectively in high-throughput proteomics.

3.2. Reproducibility

To test the reproducibility of acoustic system in rapid digestion, 3 sets of 6 standard proteins mixture (BSA, β -lactoglobulin, Myoglobin, Transferrin, β -casein, α -amylase) were processed separately at the same digested conditions by Covaris at intensity 2 in 30 minutes, cycle/burst: 200, duty cycle: 10%. Commercial search engine Thermo Proteome Discoverer (version 1.1.0.263) was used in MS/MS spectra analysis. The results showed the high reproducibility of proteolysis processed by acoustic technology with similar number of total unique peptides and number of unique peptides obtained from 3 different times of experiments (Fig. 5). High percent of sequence coverage for each protein was also achieved, especially much more than 60% with BSA, β -lactoglobulin,

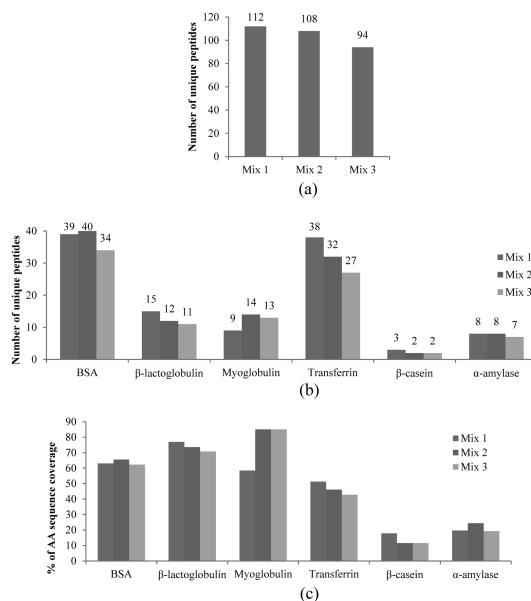


Fig. 5. The reproducibility of acoustic method in rapid digestion. 3 sets of 6 standard proteins mixture were digested by Covaris system at the same conditions: Duty cycle: 10%, Time: 30 min, Cycle/burst: 200, intensity: 2. MS/MS data were searched by Thermo Proteome Discoverer 1.1.0.263 program. (a) The number of unique peptides. (b) The number of unique peptides observed by each standard protein. (c) Percent of amino acid sequence coverage of each standard protein.

Myoglobin (Fig. 5c). Considering in details, the results showed a little difference in the number of identified peptides for each standard protein in three mixtures but their sequence coverage was approximately the same. It can be explained by miss-cleavages in sample 3, but it was not involved in protein identification.

3.3. Application for real sample

The initial experiments showed rapid and effective protein digestion could be achieved in a Covaris system, the next step was to examine direct coupling with a high-throughput proteomics platform on a real-complex protein mixture. Breast cancer cell line HS578T was reduced and alkylated, then digested by incubation with trypsin in 16 hours as usual or using acoustic technology in 30 min.

The chromatograms in Fig. 6 showed very clear similar peak patterns of peptides provided from two different digestion conditions. Moreover, in the case of 30 min with Covaris system, X!Tandem database searching identified larger amount of peptides and proteins as shown in Fig. 7. It was estimated that complete digestion was achieved by super acoustic

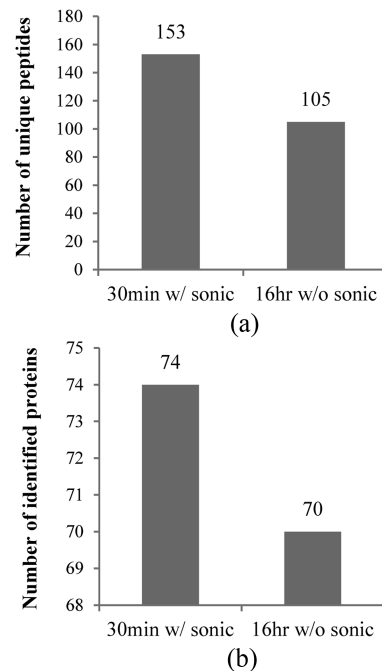


Fig. 7. Data analysis of HS578T at 2 different digestion processing using X!Tandem search engine. Samples were digested with Covaris instrument in 30 min, intensity 0.5 and 16 hr incubation at 37 °C respectively. (a) Number of identified peptides. (b) Number of identified proteins.

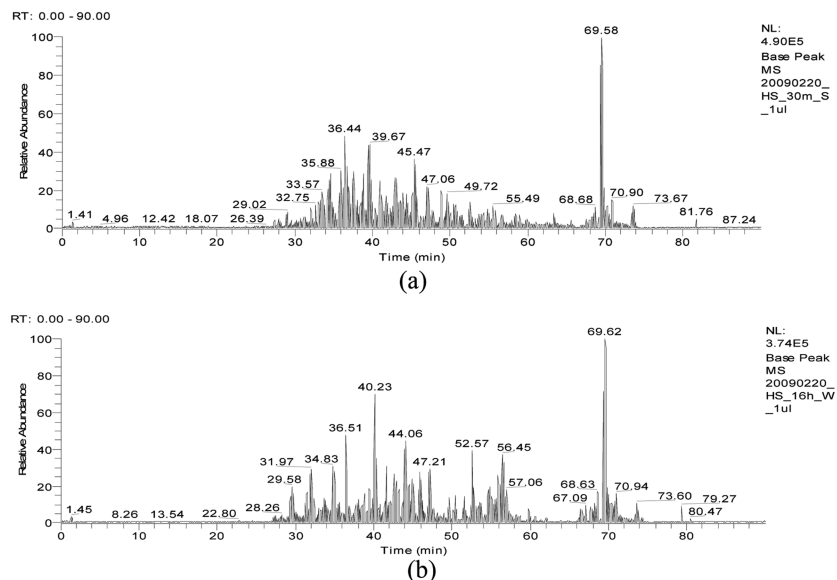


Fig. 6. Base peak chromatograms of digestion products from HS578T cancer cell line obtained at different digestion conditions. Samples were digested with Covaris instrument in intensity 2 and 30 min (a), and 16 hours (b) incubation at 37 °C respectively.

method, providing unequivocal protein identification for high-throughput proteomics analysis of complex protein mixtures.

4. Conclusion

In this study, we applied an acoustic technology for rapid protein digestion by using Covaris system and optimized operating method and parameters. The controllable wavelengths accelerate proteolysis, which not only eliminates the need for extended incubation time but also increases the digestion efficiency. In addition, by optimization time and intensity condition for acoustic, it enabled protein mixtures to be fragmented in 30 minutes, intensity 2 with higher sequence coverage, than standard protocol. For large scale clinical proteomics research, we also demonstrated rapid method worked well, but 30 times faster. This acoustic system we present here simplified the samples preparation procedure and will be helpful for the application of high-throughput analysis.

Acknowledgement

This work was supported by the Converging Research Center Program (2011K000884), by the World Class University Project (R32-2011-000102150), and by the National Research Foundation (2011K0002217, 2011K0021054) through the Ministry of Education, Science and Technology, Korea.

References

1. R. Aebersold and D. R. Goodlett, *Chem. Rev.*, **101**, 269-295 (2001).
2. R. Aebersold and M. Mann, *Nature*, **422**, 198-207 (2003).
3. T. J. Griffin, D. R. Goodlett, R. Aebersold and C. Opitz, *Biotech.*, **12**, 607-612 (2001).
4. A.-J. Moulay A. and Y. J. Xu, *J. Zhejiang Univ. Science B*, **7**(6), 411 (2006).
5. J. M. Gilmore and M. P. Washburn, *J. Proteomics*, **73**(11), 2078-2091 (2010).
6. J. A. Loo, C.G. Edmonds and R. D. Smith, *Anal. Chem.*, **63**, 2488-2499 (1991).
7. J. R. III Yates, *J. Mass Spectrometry*, **33**, 1-19 (1998).
8. J. P. Chang, D. E. Kiehl and A. Kennington, *Rapid Commun. Mass Spectrom.*, **11**, 1266-1270 (1997).
9. G. W. Slysz and D. C. Schriemer, *Anal. Chem.*, **77**(6), 1572-1579 (2005).
10. I. M. Lazar, R. S. Ramsey and J. M. Ramsey, *Anal. Chem.*, **73**(8), 1733-1739 (2001).
11. W. Sun, S. Gao, L. Wang, Y. Chen, S. Wu, X. Wang, D. Zheng and Y. Gao, *Mol. Cel. Proteomics*, **5**, 769-776 (2006).
12. G. Yao, C. Deng, X. Zhang and P. Yang, *Angew. Chem. Int. Ed.*, **49**, 8185 (2010).
13. B. E. Slentz, N. A. Penner and F. E. Regnier, *J. Chromatogr. A*, **984**(1), 97-107 (2003).
14. E. Bonneil, M. Mercier and K. C. Waldron, *Anal. Chim. Acta*, **404**, 29-45 (2000).
15. L. J. Jin, J. Ferrance, J. C. Sanders and J. P. Landers, *Lab. Chip*, **3**, 11-18 (2003).
16. K. Yamada, T. Nakasone, R. Nagano and M. Hirata, *J. Appl. Polym. Sci.*, **89**, 3574-3581 (2003).
17. K. Sakai-Kato, M. Kato and T. Toyooka, *Anal. Chem.*, **74**, 2943 (2002).
18. J. Gao, J. Xu, L. E. Locascio and C.S. Lee, *Anal. Chem.*, **73**, 2648-2655 (2001).
19. D. S. Peterson, T. Rohr, F. Svec and J. M. J. Frechet, *Anal. Chem.*, **74**, 4081-4088 (2002).
20. G. W. Slysz and D. C. Schriemer, *Rapid Commun Mass Spectrom.*, **17**, 1044-1050 (2003).
21. http://www.covarisinc.com/how_it_works.html