벼의 차세대 단백질체 분석을 위한 질량분석기 호환의 광분해성 계면활성제의 적용

신혜원 · 응웬반쯔영 · 정주용 · 이기현 · 장정우 · 윤진미 · 라비굽타 · 김선태 · 민철우

Application of mass-spectrometry compatible photocleavable surfactant for next-generation proteomics using rice leaves

Hye Won Shin · Truong Van Nguyen · Ju Young Jung · Gi Hyun Lee · Jeong Woo Jang · Jinmi Yoon · Ravi Gupta · Sun Tae Kim · Cheol Woo Min

Received: 24 September 2021 / Revised: 29 September 2021 / Accepted: 29 September 2021
© Korean Society for Plant Biotechnology

Abstract The solubilization of isolated proteins into the adequate buffer containing of surfactants is primary step for proteomic analysis. Particularly, sodium dodecyl sulfate (SDS) is the most widely used surfactant, however, it is not compatible with mass spectrometry (MS). Therefore, it must be removed prior to MS analysis through rigorous washing, which eventually results in inevitable protein loss. Recently, photocleavable surfactant, 4-hexylphenylazosulfonate (Azo), was reported which can be easily degraded by UV irradiation and is compatible with MS during proteomic approach using animal tissues. In this study, we employed comparative label-free proteomic analysis for evaluating the solubilization efficacies of the Azo and SDS surfactants using rice leave proteins. This approach led to identification of 3,365 proteins of which 682 proteins were determined as significantly modulated. Further, according to the subcellular localization prediction in SDS and Azo, proteins localized in the chloroplast were the major organelle accounting for 64% of the total organelle in the SDS sample, while only 37.5% of organelle proteins solubilized in the Azo were predicted to be localized in chloroplast. Taken together, this study

validates the efficient solubilization of total protein isolated from plant material for bottom-up proteomics. Azo surfactant is suitable as substitute of SDS and promising for bottom-up proteomics as it facilitates robust protein extraction, rapid washing step during enzymatic digestion, and MS analysis.

Keywords Mass spectrometry, Label-free quantitative analysis, Azo, SDS, MaxQuant, Perseus

서 언

벼(*Oryza sativa*, L)는 유전체가 완전히 해독되어 오믹스 분야에서 많은 연구가 수행되어 오고 있다(Darje 2013). 특히 단백질체학은 단백질에 대한 대규모 연구로 유전체학의 한계를 넘어 새로운 수준의 생명현상을 이해할 수 있으며 이를 바탕으로 단백질체와 관련된 여러 연구가 활발하게 수행되고 있다(Cox and Mann 2007). 이러한 단백질체 분석의 주요 방법으로 과거에는 주로 이차원전기영동법을 사용하여 단백질의 발현을 정량 분석하였으나 대량의 단백질을 동시에 정량화 하는 데에는 한계가 있었다(Min et al. 2019). 따라서 최근에는 고성능 질량분석기(mass spectrometry, MS)를 기반으로 하여 단백질의 발현 정도를 확인이 가능한 비표지 정량법이나 tandem mass tags (TMT) 등을 활용한 표지 정량법과 같은 기술들이 개발되어 활용되고 있다(Meng et al. 2019; Min et al. 2020a).

MS를 기반으로 한 차세대 단백질체 분석에는 단백질 그대로를 분석하는 top-down 방식과 단백질을 분해하여 펩타이드를 분석하는 bottom-up 방식 두 가지가 있다(Brown et al. 2020). Top-down 방식은 온전한 단백질 그대로를 특성화 하는데 사용하기 때문에 번역 후 변형(post-translational modifications,

H. W. Shin · T. V. Nguyen · J. Y. Jung · G. H. Lee · J. W. Jang · J. Yoon · S. T. Kim (✉) · C. W. Min (✉)
부산대학교 식물생명과학과
(Department of Plant Bioscience, Life and Industry Convergence Research Institute, Pusan National University, Miryang 50463, South Korea)
e-mail: stkim71@pusan.ac.kr, min0685@pusan.ac.kr

R. Gupta
국민대학교 교양대학
(College of General Education, Kookmin University, Seoul 02707, South Korea)

PTMs) 및 동형 단백질을 확인하는 데 bottom-up 보다 더 유리하지만 단백질 분류, 단백질 이온화 및 기체상에서의 단편화의 어려움이 있다는 한계가 있다(Zhang et al. 2013). 반면에 bottom-up 방식은 PTMs 및 서열 변형을 확인하는데 있어서는 top-down 방식에 비해 불리하지만 MS 분석에 있어 펩타이드 상태로 분석하기 때문에 단백질 상태일 때 보다 더 적합한 질량범위와 단편화 특성 등 여러 장점을 가지고 있어 더 대중적으로 사용되어 오고 있는 방식이다(Gillet et al. 2016). 두 방식 모두에서 전사체 수준의 단백질체 분석의 효율을 극대화하기 위한 출발점은 단백질을 얼마나 효과적으로 가용화(solubilization)해내는 가이다.

단백질의 가용화 효율을 높이기 위해서 대부분의 추출 용액에는 sodium dodecyl sulfate (SDS) 등의 다양한 계면활성제가 포함되는 경우가 많고 이는 질량분석에서 펩타이드의 검출 신호를 강력하게 억제하여 분석 전에 반드시 제거되어야 한다(Kalipatnapu and Chattopadhyay 2005; Loo et al. 1994). 효과적인 제거를 위해 액체 크로마토그래피, affinity spin column, filter-aided sample preparation (FASP) 방법 등이 개발되었지만 이러한 단계는 소량의 샘플일 시 필연적으로 샘플의 손실을 유발할 수밖에 없다는 단점을 가지고 있다(Hengel et al. 2012; Wiśniewski et al. 2009). 이러한 단점을 해결하고자 최근에는 광분해성 계면활성제인 4-hexylphenylazosulfonate (Azo)의 개발되었다(Brown et al. 2019). Azo 계면활성제는 UV 조사를 통해 쉽게 분해되어 질량분석기에서 분석이 가능하여 단백질 혹은 펩타이드들의 정확한 질량분석 결과를 얻을 수 있어 최근에 질량 분석기를 활용한 차세대 단백질체 연구에 많이 활용되고 있다.

따라서, 본 연구에서는 아직 식물 차세대 단백질체 분석에는 사용되지 않은 광분해성 계면활성제인 Azo를 활용하여 유용성을 밝히고자 하였다. 이를 위해서 벼 잎에서 분리된 단백질을 각각 SDS 및 광분해성의 Azo 계면활성제로 가용화한 후 비표지 단백질체 정량 분석을 수행하여 각 계면활성제들의 단백질 가용화 효율을 질량분석기를 이용하여 비교 분석하였다(Fig. 1).

재료 및 방법

식물체 준비

본 연구에 사용한 벼 종자(*Oryza sativa* L. cv. Dongjin)는 0.05% Spotak 용액(Bayer Crop Science, South Korea)에 넣어 28°C 조건에서 하룻밤 소독한 뒤, 증류수로 5회 세척하였다(Gupta et al. 2019). 소독한 종자는 28°C 암실 조건에서 젖은 티슈 페이퍼를 사용하여 발아시킨 뒤 Yoshida 용액으로 옮겨 16/8 시간의 명암 주기, 습도 70%, 온도 25°C로 유지된 성장 챔버에서 키웠다(Yoshida et al. 1976). 단백질체 분석을 위해 4주간 키운 벼

잎을 수확하였고 수확 즉시 -70°C 초저온 냉동고에 보관하였다.

Azo 합성

본 연구에 사용한 Azo 계면활성제의 합성은 이전에 보고되었던 방법에 따라 수행하였다(Brown et al. 2019, 2020). 플라스크에 0.857 mL의 4-n-hexylaniline (4 mM, n=4, C6)을 4.8 mL의 염산(10%) 과 8 mL의 ddH₂O에 넣은 뒤 10분간 10°C에서 교반하였다. 이후 4 mL의 아질산나트륨 (1 mM NaNO₂)을 첨가하면서 같은 온도조건에서 5분, 탄산나트륨(2 mM Na₂CO₃)을 조금씩 첨가하면서 3분간 교반했으며, 교반이 끝난 용액은 냉장 보관하여 차갑게 준비된 상태의 8 mL의 아황산나트륨(8 mM Na₂SO₃)이 담긴 플라스크에 여과한 뒤 4°C에서 보관하여 재결정화될 수 있도록 하였다. 마지막으로, 재결정화된 화합물을 정제하였고 이전 보고된 방법에 따라 ESI (electrospray ionization)-MS 및 ¹H-핵자기공명(NMR) 분석을 이용하여 Azo의 구조를 확인하였다

벼 잎 단백질 추출

추출방법 및 과정은 먼저 파우더 상태의 벼 잎 200 mg에 high density sucrose (HDS) buffer [37.5 mM hydroxyethyl piperazine-thanesulfonic acid (HEPES), pH 8.0, 37.5% (w/w 1.215 M) sucrose, 7.5% (v/v) glycerol, 15 mM ethylene-diamine-tetraacetic acid (EDTA), 15 mM ethylene-glycol-tetraacetic acid (EGTA), 1 mM dithiothreitol (DTT), 100x HaltTM protease inhibitor cocktail (Thermo Scientific, MA, USA)]을 추가하여 균질화시켜 준 뒤 4°C에서 3분간 600 g로 원심분리시켜 상등액을 모았다. 위의 과정을 2번 더 반복하여 남은 침전물에서 상등액을 추가적으로 얻었으며 이때, 원심분리 조건은 동일하되 HDS 용액은 처음의 절반, 3분의 1로 줄여가며 사용했다. 남아있을 지도 모르는 잔여물들을 제거하기 위해 얻어진 모든 상등액을 합친 뒤 4°C에서 3분간 600 g로 원심분리를 추가적으로 수행하였다. sucrose의 최종 농도를 12~13%로 낮추어 주기 위해 얻어진 투명한 상등액을 ddH₂O과 1:2의 비율로 희석하였다. 마지막으로 4°C에서 90분간 21,000 g로 원심분리한 뒤 얻어진 상등액을 0.07% (v/v) β-mercaptoethanol을 포함한 12.5% (w/v) trichloroacetic acid (TCA)/acetone을 이용하여 -20°C에서 침전시켜 벼 잎으로부터 단백질을 얻었다.

SDS 및 Azo 계면활성제를 사용한 단백질 가용화 및 FASP 방법을 이용한 고순도 단백질체 절편화

벼 잎에서 단백질을 분리 후, 이전에 보고되었던 FASP 방법을 이용하여 고순도 단백질체 절편화를 수행하였다(Min et al. 2020b; Wiśniewski et al. 2009). 먼저, TCA/acetone에 침전시킨 단백질(300 µg)을 30 µL의 denaturation buffer [4% SDS and

100 mM DTT in 100 mM tetraethylammonium tetrahydroborate (TEAB), pH 8.5]에 녹인 뒤 3분간 초음파 분쇄 처리를 진행하고 99°C에서 30분간 가열하였다. 이를 통해 변성된 단백질을 30K 스핀 필터(Merck Millipore, Darmstadt, Germany)에 로딩 후 최종 부피가 300 μ L가 되도록 urea buffer [8 M urea in 100 mM TEAB, pH 8.5]로 희석하였다. 다음으로 300 μ L의 urea buffer를 이용하여 14,000 g에서 원심분리를 3회를 진행하여 샘플 내의 SDS를 제거한 뒤, 상온의 암실 조건에서 200 μ L의 alkylation buffer [50 mM iodoacetamide (IAA), 8 M urea in 100 mM TEAB, pH 8.5]를 추가하여 1시간 동안 시스테인 알킬화를 진행했다. 그 후 단백질을 5% acetonitrile (ACN)이 포함된 50 mM TEAB buffer에 용해된 trypsin (enzyme-to-substrate ratio [w/w] of 1:100)으로 37°C에서 16시간동안 보관하였다. Trypsin으로 절편화된 펩타이드는 원심분리를 통해 수집하였다.

또한, Azo 계면활성제를 사용한 단백질 가용화 및 절편화 절차의 경우, TCA/acetone에 침전된 단백질을 Azo 계면활성제를 포함한 denaturation buffer [1% Azo and 100 mM DTT in 100 mM TEAB, pH 8.5]로 용해시켰다. 샘플을 3분간 초음파 분쇄 처리를 진행하고 99°C에서 30분간 가열한 후, 변성된 단백질을 30K 스핀 필터(Merck Millipore, Darmstadt, Germany)에 로딩하고 최종 부피가 300 μ L가 되도록 urea buffer [8 M urea in 100 mM TEAB, pH 8.5]로 희석하였다. 해당 방법의 경우 이전에 보고된 것에 따라서 100W 고압 수온램프(Nikon housing with Nikon HB-10101AF power supply; Nikon, Tokyo, Japan)를 사용하여 자외-가시광선(UV-vis)조사를 진행하여 Azo 계면활성제를 분해하는 추가적인 과정을 수행하였다(Brown et al. 2019). UV의 조사 후 앞서 언급된 절차와 동일하게 alkylation 및 trypsin을 처리 후 절편화된 펩타이드를 수집하였다.

펩타이드의 전처리 및 LC-MS/MS 분석

펩타이드의 농도는 Pierce quantitative fluorometric peptide assay (Thermo Scientific, MA, USA)를 사용하여 측정했으며, liquid chromatography-tandem mass spectrometry (LC-MS/MS) 분석을 위해 펩타이드 샘플은 이전 보고된 대로 HLB OASIS column을 사용하여 제조업체의 지침에 따라 염 성분을 제거해 주었다(Min et al. 2020a; Min et al. 2020c). 탈염 과정이 끝나면 건조된 펩타이드를 200 μ L의 로딩 용액(15 mM ammonium formate, 2% ACN)으로 재구성해준 뒤, 바닥에 C18 Empore 디스크 멤브레인 (3M, Bracknell, UK)을 패킹하고 POROS 20 R2 역상 resin을 200 μ L의 노란색 tip에 채워 stage-tip을 만들어 주었다. 펩타이드 용액을 stage-tip에 로딩하기 전 tip은 100% 메탄올, 100% ACN으로 세척하고 로딩 용액으로 conditioning 해주었다. 펩타이드를 로딩하고 5, 10, 15, 20, 25, 30, 35, 40, 60, 80 및 100% ACN을 포함하는 pH 10 완충 용액들을 사용하여 3개의 분획으로 분리하였다. 마지막으로 나눠진 3개의 분획 각각을 진공 원심분리기에서 동결 건조 후 LC-MS/MS

분석을 수행하였다. 얻어진 펩타이드는 2% ACN 및 0.1% formic acid (solvent-A)에 다시 용해시켜 UHPLC Dionex UltiMate® 3000 (Thermo Fisher Scientific, MA, USA) 장비를 사용하여 역상 크로마토그래피를 진행하였다. 샘플을 수집하기 위해, UHPLC에 Acclaim PepMap 100 트랩 column (100 μ m $\times$ 2 cm, nanoViper C18, 5 μ m, 100 Å)을 장착한 후 98% 용매 A를 6분 동안 6 μ L/분의 유속으로 흘려 보내주었다. 연속적으로 샘플은 Acclaim PepMap 100 모세관 컬럼(75 μ m $\times$ 15 cm, nanoViper C18, 3 μ m, 100 Å)에서 400 nL/min의 유속으로 분리되었다. LC 분석 시 용매의 구배는 90분 동안 용매-B (100% ACN 및 0.1% formic acid)를 2%에서 35%, 10분 동안 35%에서 95%, 이어서 5분 동안 90% 용매-B에서 실행되었으며, 마지막으로 15분 동안 5%로 진행하였다. LC-MS/MS를 전자 분무 이온화 소스와 함께 사중 극자 기반 질량 분석기 QExactive™ Orbitrap 고분해능 질량 분석기(Thermo Fisher Scientific, MA, Waltham, USA)에 이전 보고된 것과 동일한 매개변수를 사용하여 Top15 방법으로 연결했다(Meng et al. 2019). 그리고 모든 단백질체 질량분석 raw 데이터는 ProteomeXchange consortium을 통해 PRIDE partner repository에 PXD025132로 등록하였다(Perez-Riverol et al. 2019).

MaxQuant와 Perseus 소프트웨어를 이용한 데이터 분석

LC-MS/MS 분석으로 얻어진 데이터는 MaxQuant 소프트웨어(version 1.6.17.0)를 사용하여 이전에 보고되었던 방식에 따라 분석하였다(Tyanova et al. 2016a). 모든 데이터는 *Oryza sativa* database (*O. sativa* pepV6, 67,393 entries)를 참조하였다. Andromeda에서 첫번째에는 20 ppm, 다음에는 4.5 ppm의 기본 전구체 질량 허용오차(default precursor mass tolerances set)를 설정하여 Label-free quantification (LFQ) 데이터를 처리하였다. LFQ 데이터는 0.5 Da 질량의 허용오차(최대 2회 trypsin 절편화를 인식하지 못하는 값)를 기반으로 검색하였다. Raw data를 분석함에 있어서 fixed modification으로서 cysteine 잔기의 carbamidomethylation을 설정하고, variable modification으로서 lysine 잔기의 acetylation 및 methionine 잔기의 oxidation을 설정하였다. 또한 통계적 허용 범위 내의 펩타이드를 식별하기 위해 reverse nonsense version을 기반으로 false discovery rate (FDR)를 1%로 설정하였다.

LFQ 데이터의 multiple sample test (one-way ANOVA, Benjamini-Hochberg FDR < 0.05, fold change (FC) > 1.25), hierarchical clustering analysis (HCA), 및 principle component analysis (PCA) 등의 통계 분석은 Perseus 소프트웨어 (ver. 1.6.14.0)를 사용하여 수행하였다(Tyanova et al. 2016b). 또한 통계적으로 유의미한 차이를 보이는 단백질들의 기능 및 세포 내 위치 확인을 위하여 각각 AriGO v2.0를 활용한 Gene Ontology (GO) enrichment분석(Tian et al. 2017), CELLO2GO를 활용한 세포 내 위치 확인(Yu et al. 2014), 및 Kyoto Encyclopedia of Genes

and Genomes (KEGG) 경로 분석들을 수행하였다(Kanehisa et al. 2017).

결 과

비표지 단백질체 정량 분석

실험에 사용한 Azo는 이전에 보고된 내용(Brown et al. 2019)에 따라 합성하였고 이후 벼 잎 단백질을 Azo 및 SDS가 포함된 buffer에 녹여 준 뒤 비표지 단백질체 분석법을 활용한 비교 분석을 수행하였다(Fig. 1). 이를 통해 3,365개의 단백질

체 그룹과 이와 일치하는 22,672개의 peptide 및 21,338개의 unique peptide를 확인하였다(Fig. 2A, B). Azo 및 SDS 각각의 3 반복 샘플의 CV값은 각각 1.0% 및 0.5% 미만이었다. 또한 잠재적 오염인자 및 식별된 unique peptide의 개수가 1개 미만인 단백질과 반복 간에 확인된 intensity 값이 70% 미만인 단백질을 제거하여 총 2,343개의 단백질을 최종적으로 식별하였다(Fig. 2B). 또한, 반복 분석 간의 Pearson 상관관계 분석 결과는 각각 Azo 샘플에서는 0.96에서 0.99 사이의 값을 SDS 샘플에서는 0.99 값으로 반복 간에 높은 상관관계를 보여주었다(Fig. 2C). PCA 분석에서 PC1(93.5%)의 경우 서로 다른 계면활성제에 가용화된 두 샘플 간의 명확한 차이를 보여주

Fig. 1 Schematic workflow for bottom-up proteomic approach using rice leaves protein dissolved in the two different surfactants. Overall workflow showing the strategies for protein solubilization in the two different surfactants, trypsin digestion, and LC-MS/MS analysis for identification of differentially modulated proteins

Fig. 2 Label-free quantitative proteomic analysis of rice leaves protein solubilized in Azo and SDS surfactants, respectively. (A) Table showing the information of identified proteins in two different sample sets with replicates. (B) The narrow-down approach for the identification of significantly modulated proteins. (C) Multi-scatter plots represent Pearson's correlation coefficient values of each sample solubilized in Azo and SDS. (D) Principle component analysis of significantly modulated proteins

Fig. 3 Quantification of significantly modulated proteins of rice leaves. Volcano plot (A) and hierarchical clustering analysis (B and C) highlighting the 682 significantly modulated grouped into two major clusters based on their abundance patterns

고 있으며, PC2(6.5%)의 경우 동일한 샘플의 반복 간의 차이를 보여주고 있다(Fig. 2D). 최종적으로, multiple sample test (One-way ANOVA, Benjamini-Hochberg FDR < 0.05, FC > 1.25) 분석 결과 총 682개의 통계적으로 유의한 발현 차이를 보이는 단백질을 확인하였다(Fig. 3A, B, C 및 Table S1).

단백질 기능분석

Volcano plot 및 HCA 분석결과 통계적으로 유의한 차이를 보이는 단백질들은 2개의 주요 cluster로 분류되었다(Fig. 3A, B, C). 각 cluster의 경우 Azo에 가용화된 샘플에서 증가되는 332개 단백질(cluster_1) 과 SDS에 가용화된 샘플에서 증가되는 350개 단백질(cluster_2)인 것으로 확인되었다(Fig. 3B, C 및 Table S1). 해당 단백질들의 기능 분석을 위하여 GO 분석을 수행하였으며(Tian et al. 2017), 그 결과 biological process, molecular function, 및 cellular component에 해당하는 다양한 GO term들을 확인하였다(Fig. 4A, B 및 Table S2). 특히, cellular component 중 intracellular(각각 cluster_1과 _2에서 확인된 220개 및 279개 단백질)와 cytosol(각각 cluster_1과 _2에서 확인된 254개 및 270개 단백질)은 cellular component 범주에서 Azo 및 SDS 샘플 모두에서 특징적으로 농축되어 나타난 GO term들임을 확인하였다(Fig. 4A, B 및 Table S2). 반면, organelle의 경우 SDS에 가용화된 샘플에서는 237개의 단백질이 포함되었고, Azo에 가용화된 샘플에서는 101개의 단백질이 해당 범주의 GO term에 포함되는 것을 확인하였다(Fig. 4A, B 및 Table S2).

세포 내 단백질체의 위치 확인 분석 및 KEGG 경로 분석

유효한 발현차이를 보이는 단백질의 세포 내 위치 확인 및

KEGG 경로 분석을 수행하였다. 그 결과는 Azo (cluster_1)에서 유의미하게 증가한 332개의 단백질 중 136개(41.0%) 및 149개(44.9%)의 단백질이 각각 세포질 및 세포 소기관에 위치하는 것을 확인하였으며, 나머지 47개(14.2%)의 단백질들은 마이크로솜 막에 위치하는 것으로 나타났다(Fig. 4C 및 Table S3). 반면 SDS (cluster_2)에서는 각각 124개(37.3%), 153개(46.1%), 및 56개(16.6%)의 단백질들이 각각 세포질, 세포 소기관, 및 마이크로솜 막에 위치하는 것을 확인하였다(Fig. 4D). 각각 Azo와 SDS 계면활성제에 가용화된 단백질들의 세포 내 위치 분석의 경우 유사한 결과를 보여주고 있지만, 각각의 세포 소기관 구성요소의 비율이 다를 확인하였다. Azo에서는 엽록체(37.5%), 핵 32.9%, 미토콘드리아(15.8%), 액포(6.6%), 소포체(3.9%), 세포골격(2.6%), 퍼옥시좀(0.7%)의 순으로 나타난 반면 SDS에서는 엽록체가 64.0%로 큰 비율을 차지하였으며 미토콘드리아(16.2%), 핵(14.2%), 소포체(2.0%) 순으로 나타났고 나머지 세포소기관들이 3.5%를 차지하는 것을 추가적으로 확인하였다(Fig. 4C, D 및 Table S3).

또한, KEGG brite 데이터베이스를 기반으로 KEGG 경로 분석을 수행하였고, 그 결과 Azo에 가용화한 단백질 샘플의 경우 SDS에 가용화한 단백질 샘플에 비해 많은 수의 단백질들이 다양한 KEGG pathway에 annotation되는 것을 확인하였다(Fig. 4E)(Kanehisa et al. 2017; Kanehisa and Sato 2020). 특히, Azo 샘플의 경우 mRNA 생합성, 막 수송(membrane trafficking) 스플라이소솜, 펩티다아제 억제제 등과 관련된 단백질이 더 많이 검출되었다(Fig. 4E).

고 찰

단백질체 분석에는 다양한 종류의 계면활성제들이 선택적으로 사용되고 있으며, SDS, hexadecyltrimethylammonium

Fig. 4 Functional characterization of the significantly modulated proteins. (A, B) Cellular component in GO analysis reveals that the organelle, cytoplasm, and intracellular were found to be most enriched terms in both the samples. Subcellular localization analysis of the increased abundance proteins in Azo (cluster_1) (C) and SDS (cluster_2) (D), respectively, were carried out using UniProt and CELLO2GO web-based database. (E) The KEGG pathway analysis of significantly modulated proteins by the KEGG brite database

bromide (CTAB)와 같은 이온성과 Tritone, Tween 20과 같은 비이온성, CHAPS와 같은 양성 이온성으로 구분되며 critical micelle concentration (CMC) 등을 고려하여 선택적으로 사용된다(Kalipatnapu and Chattopadhyay 2005). 일반적으로 이온성 계면활성제들은 단백질의 구조를 완전히 변성시키는 강력한 계면활성제로 분류되며 특히 이온성 계면활성제 중 하나인 SDS는 단백질을 매우 효과적으로 변성하고 가용화할 수 있어 단백질체 분석 전 단백질 가용화 단계에서 가장 많이 사용되고 있다(Bhuyan 2010). 그러나 SDS는 chromatography의 분리를 방해하고 MS 분석 시 펩타이드의 이온화를 억제하기 때문에 이후의 분석과정을 위해서 반드시 제거되어야 하며, 이를 위해서 여러 차례의 세척 단계를 거쳐야 하고 결국 단백질의 소실과 저하를 유발하게 된다(Alfonso-Garrido et al. 2015).

최근 연구에서는 MS 분석에 적합한 계면 활성제들을 개발 및 적용하기 위한 많은 노력이 있었다. 특히 Chang et al. (2015)의 지난 연구에서는 계면 활성제와 유사한 화합물들의 스크리닝을 통해 동물 조직의 단백질체학 분석 시 SDS에 필적하는 단백질 가용화 효율을 가지고 있으며 MS 분석과 호환되는 MS-compatible degradable surfactant (MaSDeS)를 개발하였다(Chang et al. 2015). 동물의 조직에서 단백질체를 추출하기 위한 계면활성제로 MaSDeS를 사용했을 때 동물의

심장, 간 및 폐의 조직을 포함하여 일반적으로 사용되는 조직의 총 단백질 식별 수가 크게 향상되는 결과를 보여주었다(Chang et al. 2015). 또한 MaSDeS를 사용한 심장 조직 추출물에서 확인된 226개의 막 단백질이 심장 기능의 조절에 관여하는 상호 작용체를 형성함을 통해 기존의 계면 활성제를 사용하였더라면 감지되지 않았을 중요한 막 단백질을 식별할 수 있음을 시사하였다(Chang et al. 2015). 그러나 이는 산성에 불안정한 계면활성제로 절편화 이후 높은 산성 조건에서 분해되며 이러한 분해는 중요한 번역 후 변형의 손실을 유발할 수 있었다(Zhang 2015). 따라서, 이를 개선하고자 동일한 연구 그룹에서 Azo를 그 대안으로 제시했으며, 산성에서 불안정한 계면활성제인 RapiGestTM (RG)과의 직접적인 비교를 수행하였다(Yu et al. 2003; Zhang 2015). 그 결과, RG는 시료의 단백질 및 펩타이드들의 MS 신호를 크게 억제하는 반면 Azo는 MS와 호환되는 것을 입증하였다(Yu et al. 2003; Zhang 2015). 또한 동물 조직에서 Azo는 재현성 있게 단백질을 효과적으로 추출 및 가용화할 수 있고, 신속한 효소 분해 및 추가적인 탈염 단계 없이 MS 분석이 가능했으며, top-down 및 bottom-up 방식 모두에서 적용 가능한 MS 호환 계면활성제로서의 기능을 입증하여 광범위하고 높은 처리량의 단백질체 분석 시에 간소화된 전략을 제공할 수 있음을 보여주었다(Brown et al. 2020).

본 연구에서는 광분해 계면활성제인 Azo의 식물단백질체 적용 가능성 여부 확인 및 효율성 검증을 위하여 각각 SDS와 Azo를 이용하여 벼 잎 단백질체를 가용화하여 비교하는 작업을 수행하였다. 비표지 단백질체 분석 결과에서 확인한 단백질체들의 세포 내 위치를 추가적으로 검증하는 작업을 수행하였으며, 그 결과 흥미롭게도 두 샘플 사이에서 식별된 세포 소기관에 위치하는 단백질의 개수 자체는 비슷하였으나 그 구성에서 큰 차이를 보였다. SDS에 가용화된 샘플에서는 전체 세포 소기관의 구성 중 엽록체의 비율 64%로 현저하게 높았으나 상대적으로 Azo에서는 엽록체의 구성 비율이 37.5%로 더 낮고 핵과 액포, 소포체의 구성 비율이 각각 약 2배(SDS-14.2%, Azo-32.9%), 4배(SDS-1.5%, Azo-6.6%), 2배(SDS-2.0%, Azo-3.9%)더 높은 것으로 확인되었다. 또한 KEGG 경로 분석 결과에서 막 수송 관련 단백질들이 SDS에 비하여 Azo에서 높은 발현을 보여주었다. 이처럼 Azo에 가용화할 경우에는 식물체 내에서 큰 비율을 차지하는 엽록체 대비 식물 세포내에서 비교적 적은 비율을 차지하는 다양한 세포 소기관들에 위치하는 미량 단백질체들을 확인할 수 있었으며, 이는 해당 단백질체들의 가용화 정도가 SDS 대비 Azo에서 월등하게 높음을 시사한다. 또한, 최근의 연구에서는 Azo를 활용하였을 때 벼 잎 유래의 마이크로솜 막 단백질체의 가용화가 SDS보다 훨씬 더 효과적임을 입증한 바 있다. 따라서, 본 연구에서는 MS 기기와 호환성을 가지는 Azo 계면활성제를 사용한 접근 방식은 보다 간단하고 시간을 절약할 수 있다는 장점을 가지고 있으며, 기존의 SDS 계면활성제의 새로운 대체제로서 벼 잎 내의 다양한 미량 단백질체들을 가용화함에 있어서 그 효율성이 개선되는 것을 보여주었다.

적 요

식물과 동물 모두 단백질체 분석을 위해 계면활성제를 사용하여 단백질을 효과적으로 가용화 하는 것은 매우 중요하다. 여러 계면활성제 중 광분해성의 Azo는 MS 분석에 적합하며 단백질을 효과적으로 잘 가용화 하는 등의 여러 이점을 보여주었다. 그러나 대부분이 동물 단백질체 분석에 적용되어 있었으며 식물 단백질체 분석에의 적용은 미비하였다. 따라서 본 연구에서는 벼 잎의 단백질체 분석을 위한 계면활성제로 Azo를 사용하여 SDS와 비교 분석을 수행하였다. 비표지 단백질체 정량 분석, 단백질 기능 분석, 세포내 단백질체의 위치 확인 분석 및 KEGG 경로 분석을 수행하였으며, 그 결과 SDS와 비교하였을 때 Azo의 단백질 가용화 효율이 전혀 떨어지지 않음을 확인하였다. 이는 앞으로 벼 뿐만 아니라 식물단백질체 분석 시에 광분해성 계면활성제인 Azo의 적용 가능성이 무궁무진함을 의미 한다.

사 사

본 연구는 한국연구재단 중견연구자지원사업 (2019R1A2C2085868) 및 학문후속세대지원사업 (2020R1A6A3A01100427)의 지원으로 수행되었다.

References

- Alfonso-Garrido J, Garcia-Calvo E, Luque-Garcia JL (2015) Sample preparation strategies for improving the identification of membrane proteins by mass spectrometry. *Anal Bioanal Chem* 407:4893–4905
- Bhuyan AK (2010) On the mechanism of SDS-induced protein denaturation. *Biopolymers* 93:186–199
- Brown KA, Chen B, Guardado-Alvarez TM, et al (2019) A photocleavable surfactant for top-down proteomics. *Nat Methods* 16:417–420
- Brown KA, Tucholski T, Eken C, et al (2020) High-throughput proteomics enabled by a photocleavable surfactant. *Angew Chemie* 132:8484–8488
- Chang YH, Gregorich ZR, Chen, AJ, et al (2015) New mass-spectrometry-compatible degradable surfactant for tissue proteomics. *J Proteome Res* 14:1587–1599
- Cox J, Mann M (2007) Is proteomics the new genomics? *Cell* 130:95–398.
- Daric CC (2013) Post-translational modification (PTM) proteomics: Challenges and perspectives. *Mod Chem Appl* 1:1000e114
- Gillet LC, Leitner A, Aebersold R (2016) Mass spectrometry applied to bottom-up proteomics: Entering the high-throughput era for hypothesis testing. *Annu Rev Anal Chem* 9:449–472
- Gupta R, Min CW, Kim YJ, Kim ST (2019) Identification of Msp1-induced signaling components in rice leaves by integrated proteomic and phosphoproteomic analysis. *Int J Mol Sci* 20:1–17
- Hengel SM, Floyd E, Baker ES, Zhao R, Wu S, Paša-Tolić L (2012) Evaluation of SDS depletion using an affinity spin column and IMS-MS detection. *Proteomics* 12:3138–3142
- Kalipatnapu S, Chattopadhyay A (2005) Membrane protein solubilization: Recent advances and challenges in solubilization of serotonin1A receptors. *IUBMB Life* 57:505–512
- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 45:D353–D361
- Kanehisa M, Sato Y (2020) KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci* 29:28–35
- Loo RRO, Dales N, Andrews PC (1994) Surfactant effects on protein structure examined by electrospray ionization mass spectrometry. *Protein Sci* 3:1975–1983
- Meng Q, Gupta R, Min CW, et al (2019) A proteomic insight into the MSP1 and flg22 induced signaling in *Oryza sativa* leaves. *J Proteomics* 196:120–130
- Min CW, Gupta R, Agrawal GK, Rakwal R, Kim ST (2019) Concepts and strategies of soybean seed proteomics using the

- shotgun proteomics approach. *Expert Rev Proteomics* 16:795-804
- Min CW, Gupta R, Ngyuen VT, et al (2020a) A TMT-based quantitative proteomic analysis provides insights into the protein changes in the seeds of high- And low- protein content soybean cultivars. *J Plant Biotechnol* 47:209-217
- Min CW, Hyeon H, Gupta R, et al (2020b) Integrated proteomics and metabolomics analysis highlights correlative metabolite-protein networks in soybean seeds subjected to warm-water soaking. *J Agric Food Chem* 68:8057-8067
- Min CW, Park J, Bae JW, et al (2020c) In-depth investigation of low-abundance proteins in matured and filling stages seeds of *Glycine max* employing a combination of protamine sulfate precipitation and TMT-based quantitative proteomic analysis. *Cells* 9:1517
- Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, et al (2019) The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res* 47:D442-D450
- Tian T, Liu Y, Yan H, et al (2017) AgriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res* 45:W122-W129
- Tyanova S, Temu T, Cox J (2016a) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11:2301-2319
- Tyanova S, Temu T, Sinitcyn P, et al. (2016b) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13:731-740
- Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6:359-362
- Yoshida SD, Forno JC, Cock J, Gomez KA (1976) Routine procedure for growing rice plants in culture solution. In: Laboratory manual for physiological studies of rice. The International Rice Research Institute, Manila, 61
- Yu CS, Cheng CW, Su WC, et al (2014) CELLO2GO: A web server for protein subCELLular lOcalization prediction with functional gene ontology annotation. *PLoS One* 9:e99368
- Yu YQ, Gilar M, Lee PJ, Bouvier ESP, Gebler JC (2003) Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. *Anal Chem* 75:6023-6028
- Zhang X (2015) Less is more: Membrane protein digestion beyond urea-trypsin solution for next-level proteomics. *Mol Cell Proteomics* 14:2441-2453
- Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR (2013) Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* 113: 2343-2394