高等学校化学学报
CHEMICAL JOURNAL OF CHINESE UNIVERSITIES
研究论文

深度覆盖蛋白质组学质谱分析：细胞蛋白提取方法的评估

许　霞[1]，秦伟达[1]，李若萌[1]，王倩倩[2]，刘　宁[2]，李功玉[1]
（1. 南开大学化学学院分析科学研究中心，天津市生物传感与分子识别重点实验室，天津 300071；
2. 南开大学药物化学生物学国家重点实验室，天津 300350）

摘要　综合评估了分别基于尿素(Urea)、十二烷基硫酸钠(SDS)、阴离子表面活性剂(BT)和新型总RNA抽提试剂(Trizol)的蛋白质提取方法，旨在优化基于质谱的蛋白质组学的样品制备流程. 以HeLa细胞为例，利用与质谱兼容的表面活性剂BT可显著缩短提取蛋白质的总时间，减少样品制备过程中蛋白质的损失. 整合了4种蛋白质提取方法，在不依赖柱前分馏技术的前提下，从HeLa细胞中鉴定出超过7000个蛋白质；并采用无标定量的方法定量测定了其中2990个蛋白质. 值得注意的是，BT法和SDS法在提取膜蛋白方面具有更高的效率，而Urea法和Trizol法在提取细胞核和细胞质组分方面更有效. 研究结果为深度覆盖蛋白质组学提供了蛋白质提取的新型解决方案，尤其在细胞蛋白提取方面，通过整合质谱兼容型表面活性剂与传统提取方法，有效提升了蛋白质鉴定数.

关键词　表面活性剂；蛋白提取；蛋白质组学；质谱
中图分类号　O651　　**文献标志码**　A　　doi：10.7503/cjcu20240344

Mass Spectrometry-based Deep Coverage Proteome: Evaluation of Cellular Protein Extraction Methods

XU Xia[1], QIN Weida[1], LI Ruomeng[1], WANG Qianqian[2], LIU Ning[2], LI Gongyu[1*]
(*1. Tianjin Key Laboratory of Biosensing and Molecular Recognition, Research Center for Analytical Sciences, College of Chemistry, Nankai University, Tianjin* 300071, *China*;
2. State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300350, *China*)

Abstract　The current study comprehensively evaluates four different protein extraction methods based on urea, sodium dodecyl sulfate(SDS), anionic surfactants(BT), and total RNA extractor(Trizol), aiming to optimize the sample preparation workflow for mass spectrometry-based proteomics. Using HeLa cells as an example, we found that the method employing the mass spectrometry-compatible surfactant BT reagent significantly reduces the total time consumed for protein extraction and minimizes protein losses during the sample preparation process. Further integrating the four protein extraction methods, we identified over 7000 proteins from HeLa cells without relying on pre-fractionation techniques, and 2990 of them were quantified using label-free quantification. It is worth noting that the BT and SDS methods demonstrate higher efficiency in extracting membrane proteins, while the Urea and Trizol methods are more effective in extracting proteins from nuclear and cytoplasmic fractions. In summary, this study provides a novel solution for deep proteome coverage, particularly in the context of cellular protein extraction, by integrating mass spectrometry-compatible surfactants with traditional extraction methods to effectively enhance protein identification numbers.

收稿日期：2024-07-08. 网络首发日期：2024-08-23.
联系人简介：李功玉，男，博士，研究员，主要从事蛋白结构质谱方面的研究. E-mail：ligongyu@nankai.edu.cn
基金项目：国家自然科学基金(批准号：22104064，22293030，22293032)资助.
Supported by the by the National Natural Science Foundation of China(Nos.22104064，22293030，22293032).

Keywords Surfactant; Protein extraction; Proteomics; Mass spectrometry

1 Introduction

Protein extraction plays a fundamental role as the initial and indispensable step in proteomics analysis[1—8]. The efficacy of protein extraction has a direct impact on the subsequent results of mass spectrometry (MS) identification and quantitative analysis, ensuring comprehensive coverage and accurate representation of the proteome[9—12]. Protein extraction typically requires the use of a suitable cell lysis buffer to disrupt cells, facilitating the release of intracellular proteins. Currently, there are two primary strategies for protein extraction from complex biological systems and their conversion into peptides suitable for MS-based proteomic analysis[13—19].

One approach involves the use of chaotropic agents, *i. e.* , urea, which disrupt non-covalent interactions within and between proteins, enhance protein solubility, and facilitate protein digestion, followed by MS analysis[20]. Urea is commonly used for protein extraction and downstream proteomic studies, but has limited efficiency in extracting hydrophobic proteins. Another approach involves the use of anionic detergents for improved extraction of hydrophobic proteins, *i.e.* sodium dodecyl sulfate (SDS)[21]. However, even at low concentrations, SDS is incompatible with MS. In recent years, many methods have been developed to remove SDS. One prominent example is the Filter-Aided Sample Preparation (FASP) method initially described by Wisniewski *et al.* [13]. FASP utilizes a filter device to retain proteins while allowing for the removal of SDS, followed by on-filter enzymatic digestion. Though FASP is a significant advancement for proteomic sample preparation, it is not without limitations, notably sample loss due to retention in the filter and the potential for filter membrane rupture during repeated high-speed centrifugation steps. Compared to the traditional FASP method, the suspension trapping (S-Trap) significantly reduces sample processing time and is commonly used in phosphorylated proteomics sample preparation[22]. The solid-phase-enhanced sample-preparation (SP3) method, with its automation advantages, is competitive compared to traditional protein digestion methods[23]. An alternative approach to remove SDS is the Surfactant and Coagulant Assisted Sequential Extraction/On-Particle Digestion (SCAD) method[15,16]. SCAD involves the sequential addition of cold acetone, which precipitates proteins and facilitates the removal of SDS. This process allows for the preparation of samples amenable to proteomic analysis, potentially with a higher yield and lower risk of losing material, which is crucial when dealing with samples that contain low-abundance proteins.

To circumvent the limitations of both urea and SDS, some protocols employ a combination of chaotropic agents and detergents, or alternative detergents that are more compatible with downstream MS analysis, such as digitonin or Triton X-100 for the extraction of membrane proteins[24—28]. Additionally, various acid-cleavable surfactants, including RapiGest SF, ProteaseMax, AALS II, etc., have been reported for the preparation of different samples for shotgun proteomics[29—31]. Moreover, the development of new extraction techniques and reagents is an area of ongoing research in the field of proteomics, aiming to improve the recovery and analysis of hydrophobic and other difficult-to-extract proteins. Herein, we evaluated protein extraction methods with four different types of extracting agents, *e.g.*, BT-based (BT), Trizol-based (TZ), Urea-based and SDS-based (SDS) based on a label-free quantification proteomics workflow. Preliminary experiments with HeLa cell lines demonstrate that both SDS and BT methods exhibit a preference for membrane protein extraction. BT as acid-cleavable surfactants, in particular, shows enhanced efficiency in membrane protein enrichment compared to Urea and TZ methods, which more effectively enrich proteins from intracellular components, such as the nucleus and cytoplasm. Integrating these methodologies enables a more comprehensive analysis of protein composition and function, improving our understanding of cellular

mechanisms, disease pathology, and aiding in the discovery of new therapeutic targets and strategies.

2 Experimental

2.1 Materials and Measurements

LC-MS grade acetonitrile (ACN) and formic acid (FA) were obtained commercially from Thermo-Fisher Co., Ltd. (Pittsburgh, PA, USA). Purified water (conductivity of 18.2 MΩ·cm) was obtained from a Milli-Q® Reference System (Millipore Co., Bedford, MA, USA). The BCA assay kit and peptide quantitation analysis assay kit were purchased from Sangon Biotech Co., Ltd. (Shanghai, China). Total RNA extractor (Trizol) and tris (hydroxymethyl) amino methane (purity 99.9%) were purchased from Thermo-Fisher Co., Ltd. (Pittsburgh, PA, USA). Sodium chloride (purity 99.5%), urea (purity 99%), ammonium bicarbonate (purity 98%) and Iodoacetamide (IAA) (purity 98%), were commercially obtained from Sigma-Aldrich Co., Ltd. (St. Louis, MO, USA). Dithiothreitol (DTT) (purity 99%) was provided commercially by Beyotine Biotech Co., Ltd. (Shanghai, China). Trypsin, LysC and BT surfactants were purchased from Beijing Shengxia Proteins Scientific Co., Ltd. (Beijing, China). SDS (purity 98%) was purchased from Shanghai Macklin Biochemical Co., Ltd. (Shanghai, China). Unless otherwise, all solvents used in this study were of HPLC grade supplied by Sigma-Aldrich Co., Ltd. (St. Louis, MO, USA). Dulbecco's modified Eagle's medium (DMEM) and fetal bovine serum (FBS) were purchased from Gibco Co., Ltd. (Carlsbad, CA, USA). HeLa cell line was from the American Type Culture Collection (Manassas, VA, USA). RapiGest SF was purchased from the Waters Corporation (Milford, MA, USA).

LTQ Orbitrap Eclipse mass spectrometry (MS) coupled to a Dionex UltiMated 3000 ultra-performance liquid chromatography (UPLC) system and cell incubator (Thermo-Fisher Scientific, Waltham, MA, USA), Eppendorf concentrator plus (Eppendorf, Germany), XM-650T ultrasonic homogenizer (Xiaomei Ultrasonic Instrument Co., Ltd., China), DYY-6C gel electrophoresis system (Beijing LIUYI Biotechnology Co., Ltd. China), iCEN-24R high-speed centrifuge (Hangzhou Allsheng Instruments Co., Ltd., China).

2.2 Protein Extraction and Digestion

2.2.1 *For Urea* HeLa cells precipitation was lysed in lysis buffer containing 8 mol/L urea, 50 mmol/L tris-HCl, 5 mmol/L $CaCl_2$, 30 mmol/L NaCl and protease inhibitor (1∶100, volume ratio), and add appropriate amount of HCI to adjust *ca.* pH=8.0. Each sample was prepared in triplicate. Then use a probe sonicator (40% power) to extract proteins from cells. Perform BCA protein assay according to the manufacturer's protocols. According to the measured protein concentration, take a certain volume of supernatant containing about 100 μg of protein for subsequent experiments. Prepare 100 mmol/L DTT in tris buffer (pH=8.0). Add DTT to each sample to a final concentration of 10 mmol/L, and incubate at 37 ℃ for 30 min. Subsequently, prepare 200 mmol/L IAA in tris buffer and add IAA to each sample to a final concentration of 50 mmol/L, incubate in the dark at room temperature for 30 min. Adding DTT again to a final concentration of 10 mmol/L to react with excess IAA and incubate at room temperature for 5 min. Each sample was subjected to digestion with LysC and trypsin (1∶100, mass ratio) and incubated at 37 ℃ overnight (urea<1 mol/L). The samples were quenched with 10% (volume fraction) trifluoroacetic acid (TFA) to a final concentration of 0.25% and centrifuged at 14000*g* for 15 min. Desalting the supernatant using Sep-Pak C_{18} (Waters) according to the manufacturer's protocols and drying down the sample. Resuspend samples with an appropriate volume of 0.1% (volume fraction) FA-H_2O, centrifuge at 14000*g* for 10 min. Next, determine the peptide concentration of the supernatant using an assay kit and aliquot an appropriate number of samples for LC-MS/MS analysis.

2.2.2 *For BT Surfactant* Appropriate number of cells were dissolved in buffer solution (3% volume fraction BT, 50 mmol/L NH_4HCO_3 buffer, pH=8.0) and sonicated for 10 min. A protease inhibitor was added to the

lysis buffer in advance. The protein supernatant(*ca.* 100 μg) after centrifugation was reduced with 10 mmol/L DTT(50 mmol/L NH_4HCO_3 buffer, pH=8.0) and alkylated with 50 mmol/L IAA(50 mmol/L NH_4HCO_3 buffer, pH=8.0). Transfer the protein extract after alkylated to a 10 kDa ultrafiltration disc(UFD). The 10 kDa UFD was pre-washed with digestion buffer(0.1% volume fraction BT) before use. Mix 200 μL of digestion buffer thoroughly with the samples, and centrifuge them for 10 min at 14000*g* at 20 ℃. Add another 200 μL of digestion buffer and centrifuge at 14000*g* for 10 min at 20 ℃. Repeat one more time. Digest the protein overnight on the membrane at a ratio of 1∶100(mass ratio) using both trypsin and LysC. Add 50%(volume fraction) FA-H_2O to adjust the pH value of 2.0 to quench the reaction and incubate the mixture at 45 ℃ for 30 min for BT degradation. And then centrifuge the filter units at 14000*g* for 20 min at 20 ℃. Collect the flow-through and add 100 μL 10%(volume fraction) ACN and centrifuge the UFD units at 14000*g* for 10 min. Repeat for 1×time. Combine the flow-through and subject to peptide assay.

2.3 LC-MS/MS Analysis

Each sample was dissolved in 0.1%(volume fraction) FA-H_2O before being loaded onto a 75 μm inner diameter homemade microcapillary column that is packed with 15 cm of bridge ethylene hybrid C18 particles (17 μmol/L, 13 nm, Waters). Samples were analyzed on LTQ Orbitrap Eclipse MS coupled to a Dionex UltiMated 3000 UPLC system. A mobile phase composed of 0.1%(volume fraction) aqueous FA(A) and ACN containing 0.1%(volume fraction) FA(B) was pumped into a homemade microcapillary column with a gradient program as below: 0—25 min, 7%—15%B; 25—110 min, 15%—30%B; 110—140 min, 30%—45%B, 140—141 min, 45%—95%B, 141—150 min, 95%B; and flow rate, 300 nL/min. The injection volume was fixed at 2.0 μL. The instrument was operated in data-dependent acquisition(DDA) mode. Survey scans of peptide precursors were recorded in the orbitrap(OT) within the *m/z* 300—2000 range at a resolution of 120 K, followed by MS/MS acquisition in the linear ion trap(IT). Standards were selected as the automatic gain control(AGC) targets for MS^1 and MS^2. Precursors were selected for fragmentation for continuous 3 s with stepped normalized collision energies of 27, 30 and 33. The maximum injection time was set to be 50 and 110 ms for MS^1 and MS^2 scans, respectively. Precursors were subject to dynamic exclusion for 45 s with a δ 10 tolerance. For all data acquisition including high-field asymmetric waveform ion mobility spectrometry (FAIMS), the dispersion voltage(DV) was set at 5000 V, the compensation voltages(CVs) were set at −40, −60, and −80 V. All samples were acquired in positive mode. Each sample was equilibrated before injection.

2.4 Data Analysis

Protein identification and quantification were performed using MaxQuant[32] (Version 1.6.3.4). Raw files were searched against the Uniport Homo sapiens reviewed database (June 2022) with trypsin/P and LysC/P selected as enzymes. For MS^1 scans, a precursor ion mass tolerance of δ 10 was used, and two missed cleavages were allowed. Fragment ion tolerance was set to 0.02 Da for the Orbitrap MS^2 detections and 0.5 Da for the IT detections. The false discovery rate(FDR) was set to 1% for both protein and peptide identification by the target decoy strategy. Carbamidomethylation of cysteine residues(+57.02146 Da) was chosen as the fixed modification, and variable modification included oxidation of methionine residues(+15.99492 Da) and acetylation at protein N-termini(+42.01056 Da). The quantification method was set to MaxLFQ algorithm and protein quantitative ratios were determined using a minimum of two quantified peptides. All other parameters were set as default.

All statistical analyses were accomplished and filtered for reverse proteins, proteins only identified by site and potential contaminants by Perseus(version 1.6.14.0). Owing to large inter-individual differences in protein expression, proteins that were only detected≤2 times in one group were filtered out after grouping. The missing data points were imputed with the normal distribution method, with a width of 0.3 and a downshift of

1.8. The resulting dataset was utilized for subsequent statistical analysis and computation. The filters dataset was further analyzed and a two-sample Student's *t*-test with a two-tailed distribution and one-way ANOVA were also performed. All *p*-values were further subjected to multiple testing corrections using the Benjamini-Hochberg method. Data visualization analyses including hierarchical clustering, Gene Set Variation Analysis (GSVA), volcano plot, Pearson's correlation analysis and chord diagram were achieved using R packages and Hiplot Pro online software. For protein intensity profiling, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway were enriched using DAVID[33] bioinformatics resources with an FDR cutoff of 5%.

2.5 Data Availability

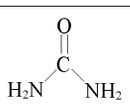
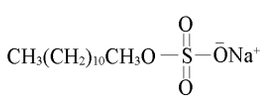
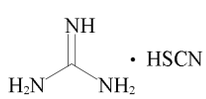
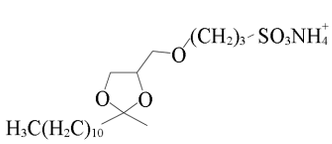
The MS proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository with the dataset identifier PXD050395.

3 Results and Discussions

3.1 BT-based Protein Extraction

The hydrophobic nature of membrane proteins and their limited water solubility often necessitate the use of surfactants to enhance solubility and stability. Traditional surfactants such as SDS can induce conformational changes in membrane proteins, adversely affecting the precision of quantitative proteomic analysis[13,21,34]. BT, as an alternative surfactant, improves membrane protein solubility and stability while preserving their native conformation, leading to heightened sensitivity and accuracy in MS detection. BT can be readily degraded under acidic conditions, streamlining the protocol by removing the need for additional isolation steps, thus conserving time and resources. This approach circumvents desalting steps, reducing the potential loss of low-abundance proteins and isoforms. Moreover, the Trizol reagent effectively segregates proteins from nucleic acids, minimizing interference from ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) with particular low-abundance and labile proteins, which enhances protein sequence coverage. Table 1 provides detailed information on different protein extraction methods.

Table 1 Information for four protein extraction methods

Abb.	Full name	Structure	Lysis buffer	Clean-up method	Time
Urea(U)	Urea-based Protein Extraction		50 mmol/L Tris buffer, 8 mol/L urea, 5 mmol/L $CaCl_2$, 30 mmol/L NaCl	C_{18} SepPak desalting	*ca.* 10 h
SDS(S)	Surfactant and Chaotropic Agent Assisted Sequential Extraction/On-Pellet Digestion		50 mmol/L Tris buffer, 4% SDS	80% Acetone precipitation+C_{18} SepPak desalting	*ca.* 22 h
Trizol(T)	Trizol, Total RNA Extractor		Trizol, chloroform	Evaporation+C_{18} SepPak desalting	*ca.* 13 h
BT(B)	BT Surfactant		50 mmol/L NH_4HCO_3, 3% BT	pH=2.3—3.0, 45 ℃, On-filter degradation, desalting-free	*ca.* 1 h

Biological research often requires the identification and separation of charged compounds from bio-samples. The anionic surfactant BT offers higher biocompatibility and reduced toxicity compared to SDS, and its superior salting-out effect enhances the separation efficiency of charged biomolecules, thus improving sample preparation. Fig. 1 depicts the BT-based proteomics workflow, highlighting the on-filter degradation pathway and the desalting-free sample preparation process. The general workflow of proteomics (urea, SDS,

trizol) is illustrated in Fig.S1 (see the Supporting Information of this paper). The BT method simplifies the removal process by requiring only heat-induced degradation, in contrast to the more complex SCAD protocol for SDS removal. The long tail of the alkyl group of the BT agent can be removed upon heating under acidic conditions, releasing the anionic glycerol part (Fig.1). Total clean-up process for BT removal can be as short as 1 h, greatly saving sample preparation time compared to Urea- and SDS-based protein extractions (Table 1). Additionally, BT's affinity for membrane proteins addresses the shortcomings of urea in extracting membrane-associated proteins. The performance of this new method was appraised through an unbiased comparison with various protein extraction techniques.

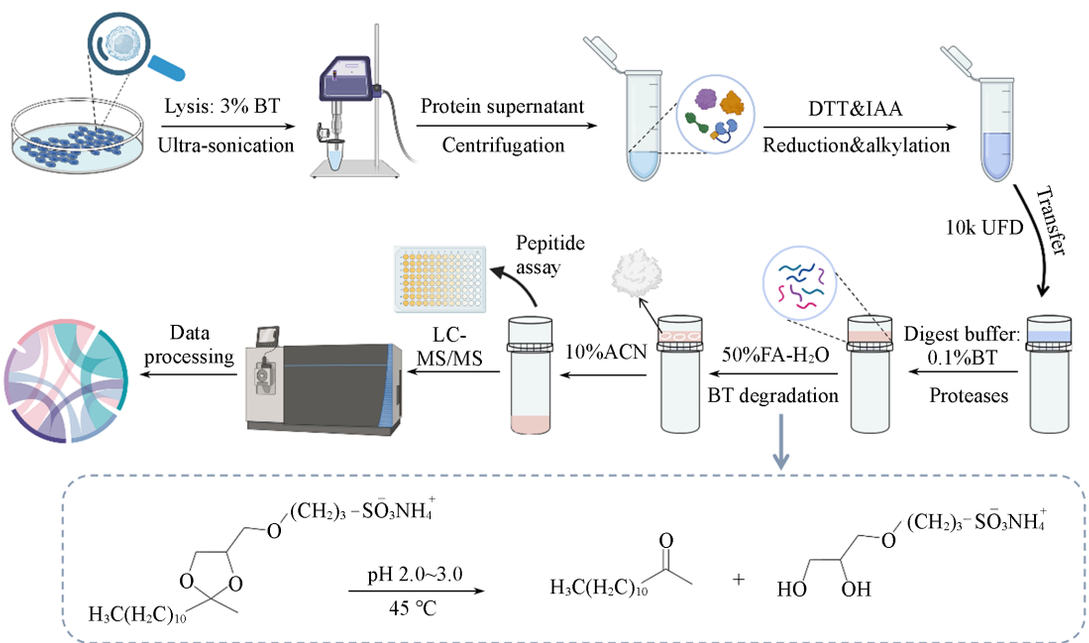


Fig. 1 BT-based proteomics workflow

Key steps involve: dissolving cell precipitates with 3% BT; transferring the extracted cellular proteins to a 10 kDa filter for digestion with a buffer of 0.1% BT; and degrading BT under acidic conditions (pH=2.0—3.0, 45 ℃). The overall clean-up time for BT-based extraction workflow is *ca.* 1 h. More details can be found in Experimental.

3.2 Protein Extraction Performance Evaluation

To assess protein extraction efficacy, three biological replicates were conducted using four methods (Urea, BT, SDS, and Trizol) on human HeLa cell lines. Qualitative analyses of extracted proteins *via* SDS-PAGE [(Fig.2(A)] demonstrated similar extraction capabilities among the methods, as evidenced by the distribution of protein bands with equal extract loading. BT's protein mass distribution was akin to that of SDS, while Trizol's distribution was like urea, aligning with their respective hydrophobicity preferences. Peptide yields from the BT method were slightly higher [*ca.* 2% increase, Fig.2(B)] compared to other methods, likely attributable to its desalting-free process, which enhances sample recovery and minimizes loss.

Proteomics data quality assessment (Fig.3 and Fig.S2, see the Supporting Information of this paper) indicates that the BT method maintains protein identification performance despite bypassing a multi-step desalting process. This is evidenced by the range of identified protein abundances spanning six orders of magnitude [Fig.2(C)] and the distribution of sequence coverage [Fig.2(D)]. A Venn diagram [Fig.2(E)] displays 4565 proteins identified by all four methods out of a total of 7005, indicating a 65% overlap in protein IDs. Comparative analysis [Fig.2(F)] demonstrates the notable efficacy of BT and SDS in identifying unique

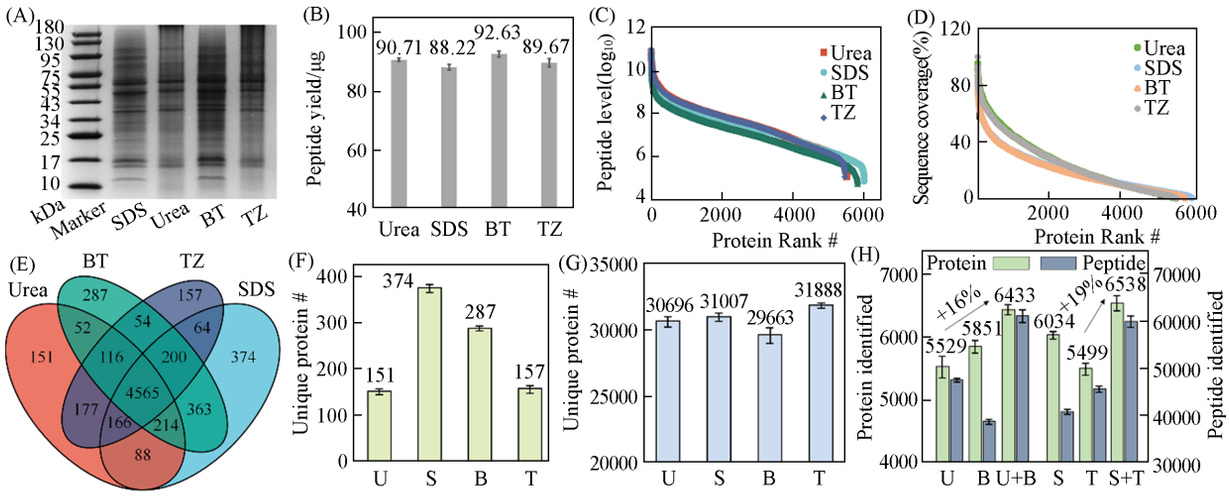


Fig. 2 Performance checking on different protein extraction methods, Urea(U), SDS(S), BT(B) and Trizol(T)

(A) Representative SDS-PAGE characterization of extracted proteins; (B) peptide yield from the same amount of starting materials; (C) protein abundance distribution; (D) protein sequence coverage; (E) Venn diagram for protein IDs; (F) unique protein ID; (G) unique peptide ID and (H) BT- and Trizol-enhanced protein identification. All data are from triplicate biological replicates with error bars showing standard deviation.

proteins, while the Trizol method yields the highest number of unique peptide identifications [Fig. 2(G)]. Consequently, integrating multiple extraction methods could yield a more comprehensive proteomic analysis.

To harness the benefits of multiple extraction methods, proteins and peptides from urea and BT were grouped together, and those from SDS and TZ were combined into another group [Fig. 2(H)]. Duplicate entries were eliminated from all groups to ensure a unique dataset. This integration significantly augmented the number of proteins and peptides detected. The combined Urea and BT group showed increases of 16% in proteins and 28% in peptides, while the BT-only group saw increases of 10% in proteins and 58% in peptides. The combined SDS and TZ group also exhibited substantial increments, with proteins increasing by 8% and 19%, and peptides by 47% and 31%, respectively, when compared to their single-method counterparts. Additional combinations also yielded notable enhancements in protein identification (Fig. S4, see the Supporting Information of this paper). Based on the performance of the four extraction reagents and the observed increases in proteins and peptides, optimal combinations for extraction are recommended as delineated in Fig.2(H).

We then conducted unique protein analysis for four protein extraction methods (Urea, SDS, BT and Trizol). Mass distribution analysis shows BT extracts similarly sized proteins as SDS, while isoelectric point analysis reveals no significant differences across the four extraction methods (Fig. S5 and Fig. S6, see the Supporting Information of this paper). However, BT extracts a significantly higher proportion of hydrophobic proteins (15.6%) compared to Urea (10%) and SDS (12.5%) (Fig.S7, see the Supporting Information of this paper). The hydrophobic protein proportion for TZ (11.8%) falls between urea and SDS.

Furthermore, we present a comparative analysis of the protein extraction methods and properties of BT and RapiGest, two MS-compatible surfactants, using HeLa cell lines. The degradation pathways (Fig. 1 and Fig.S8, see the Supporting Information of this paper) reveal that the NH_4^+ ions released upon BT degradation are more advantageous for MS compared to the Na^+ ions released by RapiGest. The compatibility of ammonia ions with MS allows BT concentrations to reach 3% (approximately 100 mmol/L NH_4^+), while 100 mmol/L Na^+ from RapiGest would cause severe ion suppression. Consequently, BT concentrations can reach 3%, whereas RapiGest is traditionally used at 0.1%. This difference in the applicable concentration range contributes to

the observed variations in protein extraction and identification.

BT demonstrates unique advantages over RapiGest in peptide yield and membrane protein extraction, with increases of 4.5% in proteins and 11.5% in peptides (Fig.S9, see the Supporting Information of this paper). Analysis of the unique proteins identified 678 and 446 proteins for BT and RapiGest, respectively. Gravy values and molecular weights of the unique proteins were calculated to analyze their physical properties. BT shows a 2.39% increase in preference for hydrophobic proteins compared to RapiGest. The cellular component comparison indicates that BT tends to enrich membrane proteins. Biological process analysis revealed that unique proteins identified by BT are more readily enriched in processes such as cancer and cellular energy metabolism. These findings provide valuable insights for further research on the roles of these proteins in biological processes, including cancer. The development of the BT method offers new opportunities for effective membrane protein extraction. Future research will focus on enhancing and refining the BT method to expand its applicability and gain precise, comprehensive insights into membrane protein functions and their associations with various diseases.

3.3 Quantitative Proteome Profiling

Following the comprehensive protein ID characterization, protein quantification analysis was conducted to discern extraction bias from various strategies. Post data quality evaluation for reproducibility (Fig.S10 and Fig.S11, see the Supporting Information of this paper), hierarchical clustering was performed on all quantified proteins to examine the profiles derived from different protein extraction methods [Fig.3(A)]. The heatmap illustrates column-based clustering of biological replicates across extraction methods. Further analysis, based on protein abundance from different extraction methods, generated four clusters. GSVA was employed to identify changes in gene or gene set expression related to functionality, with analysis conducted on cellular components, selected biological processes, and pathways for each cluster [Fig.3(B) and Fig.S12, see the Supporting Information of this paper]. The hierarchical clustering dendrogram showed no significant intragroup differences, indicating well-defined clusters for each method.

Highly extracted proteins [Fig.3(B)] obtained *via* BT and SDS methods (Clusters 1 and 4) were notably enriched with various membrane proteins. Conversely, proteins enriched *via* Urea and TZ methods (Clusters 2 and 3) were primarily localized in the cytoplasm or associated with nuclear components. A small number of cellular membrane-associated components were identified in Clusters 2 and 3. Further analysis of biological processes and KEGG pathways (Fig.S12, see the Supporting Information of this paper) confirms the preferences for each extraction method.

Pairwise comparisons between groups were facilitated by Student's t-test for all possible combinations, visualized *via* a volcano plot [Fig.3(C)]. Proteins with P value<0.05 and fold change (FC)>2 were defined as significantly differentially extracted proteins (DEPs), represented by colored dots on the plot. A comprehensive analysis was conducted on both up-regulated and down-regulated proteins, with the top ten proteins with the lowest P value highlighted. A significant disparity in the number of DEPs between the Urea and BT groups compared to the Urea and TZ groups was observed, suggesting a more pronounced distinction between Urea and BT groups and more prominent changes in protein extraction. We examined more closely two pairs of samples: Urea versus BT, which reflects differences between the traditional and novel methods. Proteins highly expressed in the Urea group were primarily associated with cellular fundamental processes such as energy metabolism and protein degradation, while those highly expressed in the BT group were mainly involved in chromatin and gene expression regulation, as well as cellular stress responses. These analytical results hold significant value for understanding differences in protein extraction with different extraction methods and their potential impact on interrogating cellular functions and metabolisms.

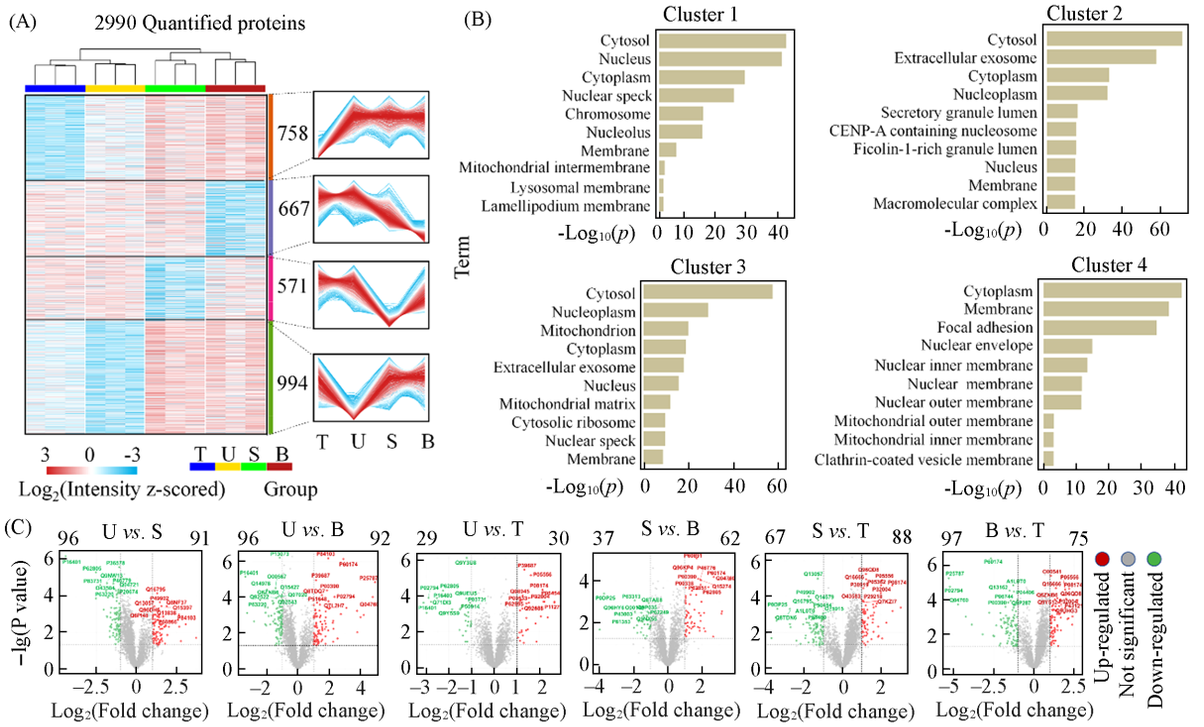


Fig. 3 Protein profiles with different extraction methods

(A) Hierarchical clustering of 2990 quantified proteins and four clusters of these proteins are depicted on different intensity profiles; (B) the number of proteins and cellular components in each cluster was marked; (C) the volcano plot illustrates the pairwise comparison of protein expression levels between different methods. Dots above the horizontal dashed line represent significantly altered proteins(two-sided t-test, P value<0.05, adjusted for multiple comparisons using the Benjamini-Hochberg method). Down-regulated proteins, indicated by green color, while upregulated proteins are represented by red color(protein fold change>2).

4 Conclusions

In this study, we propose a novel protein extraction method based on the surfactant BT. By combining the Urea, SCAD, BT, and TZ protein extraction methods, more than 7000 distinct protein groups were identified in HeLa cells. The BT method shows significant advantages in terms of peptide recovery rate and number of protein identifications. The analysis of unique proteins demonstrates the distinctiveness among the four protein extraction methods. Analysis of cellular components demonstrates that SDS and BT methods show a preference for the extraction of membrane proteins, with the BT method demonstrating greater extraction efficiency. Conversely, the Urea and TZ methods exhibit better enrichment for intracellular components, including the nucleus and cytoplasm. Through the integration of novel and traditional protein extraction methods, our research provides a more comprehensive analysis of low-abundance proteins compared to the use of traditional methods alone. This underscores the importance of selecting protein extraction methods carefully and the potential benefits that arise from combining multiple methods for conducting a more thorough analysis.

As essential components of the cell membrane, membrane proteins play critical roles in numerous biological processes. Efficient extraction of these proteins contributes to a comprehensive understanding of cellular biology and the mechanisms involved in disease development. Moreover, it provides a foundation for the development of therapeutic strategies targeting membrane proteins. The development of the BT method has introduced a novel opportunity for the effective extraction of membrane proteins. Future investigations will focus on continuous enhancement and refinement of the BT method to expand its applicability and yield precise and comprehensive insights into the study of membrane protein functionality and its association with various diseases.

支持信息见http: //www.cjcu.jlu.edu.cn/CN/10.7503/cjcu20240344.

参 考 文 献

[1] Shen B., Yi X., Sun Y. T., Bi X. J., Du J. P., Zhang C., Quan S., Zhang F. F., Sun R., Qian L. J., Ge W. G., Liu W., Liang S., Chen H., Zhang Y., Li J., Xu J. Q., He Z. B., Chen B. F., Wang J., Yan H. X., Zheng Y. F., Wang D. L., Zhu J. S., Kong Z. Q., Kang, Z. Y., Liang, X., Ding, X., Ruan G., Xiang N., Cai X., Gao H. H., Li L., Li S. N., Xiao Q., Lu T., Zhu Y., Liu H. F., Chen H. X., Guo T. N., *Cell*, **2020**, *182*(1), 59—72
[2] Li Z. H., Tremmel D. M., Ma F. F., Yu Q. Y., Ma M., Delafield D. G., Shi Y. T., Wang B., Mitchell S. A., Feeney A. K., Jain V. S., Sackett S. D., Odorico J. S., Li L. J., *Nat. Commun.*, **2021**, *12*(1), 1020
[3] Martinez-Val A., Fort K., Koenig C., Van der Hoeven L., Franciosa G., Moehring T., Ishihama Y., Chen Y. J., Makarov A., Xuan Y., Olsen J. V., *Nat. Commun.*, **2023**, *14*(1), 3599
[4] Aebersold R., Mann M., *Nature*, **2003**, *422*(6928), 198—207
[5] Andersen J. S., Wilkinson C. J., Mayor T., Mortensen P., Nigg E. A., Mann M., *Nature*, **2003**, *426*(6966), 570—574
[6] Yang L., Weng S., Qian X. H., Wang M. C., Ying W. T., *Anal. Chem.*, **2022**, *94*(25), 8827—8832
[7] Yu Q., Liu X. Y., Keller M. P., Navarrete-Perea J., Zhang T., Fu S. P., Vaites L. P., Shuken S. R., Schmid E., Keele G. R., Li J., Huttlin E. L., Rashan E. H., Simcox J., Churchill G. A., Schweppe D. K., Attie A. D., Paulo J. A., Gygi S. P., *Nat. Commun.*, **2023**, *14*(1), 555
[8] Meng Y. Q., Chen J. Y., Liu Y. Q., Zhu Y. P., Wong Y. K., Lyu H. N., Shi Q. L., Xia F., Gu L. W., Zhang X. W., Gao P., Tang H., Guo Q. Y., Qiu C., Xu C. C., He X., Zhang J. Z., Wang J. G., *J. Pharm. Anal.*, **2022**, *12*(6), 879—888
[9] Davis S., Charles P. D., He L., Mowlds P., Kessler B. M., Fischer R., *J. Proteome Res.*, **2017**, *16*(3), 1288—1299
[10] Wang W., Tai F. J., Chen S. N., *J. Sep. Sci.*, **2008**, *31*(11), 2032—2039
[11] Doellinger J., Schneider A., Hoeller M., Lasch P., *Mol. Cell. Proteomics*, **2020**, *19*(1), 209—222
[12] Zhang Y., Cai Q. H., Luo Y. X., Zhang Y., Li H. L., *J. Pharm. Anal.*, **2023**, *13*(1), 63—72
[13] Wiśniewski J. R., Zougman A., Nagaraj N., Mann M., *Nat. Methods*, **2009**, *6*(5), 359—362
[14] Washburn M. P., Wolters D., Yates J. R. 3rd, *Nat. Biotechnol.*, **2001**, *19*(3), 242—247
[15] Ma F. F., Liu F. B., Xu W., Li L. J., *J. Proteome Res.*, **2018**, *17*(8), 2744—2754
[16] Ma F. F., Tremmel D. M., Li Z. H., Lietz C. B., Sackett S. D., Odorico J. S., Li L. J., *J. Proteome Res.*, **2019**, *18*(8), 3156—3165
[17] Wiśniewski J. R., *Anal. Chim. Acta*, **2019**, *1090*, 23—30
[18] Sethi M. K., Downs M., Zaia J., *Mol. Omics*, **2020**, *16*(4), 364—376
[19] Li L., Sun C. J., Sun Y. T., Dong Z., Wu R. X., Sun X. T., Zhang H. B., Jiang W. H., Zhou Y., Cen X. F., Cai S., Xia H. G., Zhu Y., Guo T. N., Piatkevich K. D., *Nat. Commun.*, **2022**, *13*(1), 7242
[20] Grzybowski A. T., Shah R. N., Richter W. F., Ruthenburg A. J., *Nat. Protoc.*, **2019**, *14*(12), 3275—3302
[21] Serra A., Gallart-Palau X., Dutta B., Sze S. K., *J. Proteome Res.*, **2018**, *17*(7), 2390—2400
[22] HaileMariam M., Eguez R. V., Singh H., Bekele S., Ameni G., Pieper R., Yu Y. B., *J. Proteome Res.*, **2018**, *17*(9), 2917—2924
[23] Hughes C. S., Moggridge S., Müller T., Sorensen P. H., Morin G. B., Krijgsveld J., *Nat. Protoc.*, **2019**, *14*(1), 68—85
[24] Yu D. H., Wang Z., Cupp-Sutton K. A., Guo Y. T., Kou Q., Smith K., Liu X. W., Wu S., *J. Am. Soc. Mass Spectr.*, **2021**, *32*(6), 1336—1344
[25] Dimayacyac-Esleta B. R. T., Tsai C. F., Kitata R. B., Lin P. Y., Choong W. K., Lin T. D., Wang Y. T., Weng S. H., Yang P. C., Arco S. D., Sung T. Y., Chen Y. J., *Anal. Chem.*, **2015**, *87*(24), 12016—12023
[26] Chen Y. J., Roumeliotis T. I., Chang Y. H., Chen C. T., Han C. L., Lin M. H., Chen H. W., Chang G. C., Chang Y. L., Wu C. T., Lin M. W., Hsieh M. S., Wang Y. T., Chen Y. R., Jonassen I., Ghavidel F. Z., Lin, Z. S., Lin K. T., Chen C. W., Sheu P. Y., Hung C. T., Huang K. C., Yang H. C., Lin P. Y., Yen T. C., Lin Y. W., Wang J. H., Raghav L., Lin C. Y., Chen Y. S., Wu P. S., Lai C. T., Weng S. H., Su K. Y., Chang W. H., Tsai P. Y., Robles A. I., Rodriguez H., Hsiao Y. J., Chang W. H., Sung T. Y., Chen J. S., Yu S. L., Choudhary J. S., Chen H. Y., Yang P. C., Chen Y. J., *Cell*, **2020**, *182*(1), 226—244
[27] Liu Z., Wang K. Y., Ye M. L., *Anal. Chem.*, **2023**, *95*(34), 12580—12585
[28] Dang Y. M., Jiang N., Wang H., Chen X. C., Gao Y., Zhang X. Y., Qin G. X., Li Y. M., Chen R. B., *J. Proteome Res.*, **2020**, *19*(3), 1298—1309
[29] Frei A. P., Jeon O. Y., Kilcher S., Moest H., Henning L. M., Jost C., Plückthun A., Mercer J., Aebersold R., Carreira E. M., Wollscheid B., *Nat. Biotechnol.*, **2012**, *30*(10), 997—1001
[30] Ji Y. H., Liu M. J., Bachschmid M. M., Costello C. E., Lin C., *Anal. Chem.*, **2015**, *87*(11), 5500—5504
[31] Li M., Powell M. J., Razunguzwa T. T., O'Doherty G. A., *J. Org. Chem.*, **2010**, *75*(18), 6149—6153
[32] Tyanova S., Temu T., Cox J., *Nat. Protoc.*, **2016**, *11*(12), 2301—2319
[33] Huang D. W., Sherman B. T., Lempicki R. A., *Nat. Protoc.*, **2009**, *4*(1), 44—57
[34] Brown K. A., Chen B. F., Guardado-Alvarez T. M., Lin Z. Q., Hwang L., Ayaz-Guner S., Jin S., Ge Y., *Nat. Methods*, **2019**, *16*(5), 417—420

(Ed.:X, N, K)