High throughput and accurate serum proteome profiling by integrated sample preparation technology and single-run data independent mass spectrometry analysis

Lin Lin, Jiaxin Zheng, Quan Yu, Wendong Chen, Jinchun Xing, Chenxi Chen, Ruijun Tian

Materials Characterization & Preparation Center, Southern University of Science and Technology, Shenzhen 518055, China
Department of Urology and Center of Urology, The First Affiliated Hospital of Xiamen University, Xiamen 361003, China
Division of Advanced Manufacturing, Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China
Department of Chemistry and Guangdong Provincial Key Laboratory of Cell Microenvironment and Disease Research, Southern University of Science and Technology, Shenzhen 518055, China

Abstract

Mass spectrometry (MS)-based serum proteome analysis is extremely challenging due to its high complexity and dynamic range of protein abundances. Developing high throughput and accurate serum proteomic profiling approach capable of analyzing large cohorts is urgently needed for biomarker discovery. Herein, we report a streamlined workflow for fast and accurate proteomic profiling from 1 μL of blood serum. The workflow combined an integrated technique for highly sensitive and reproducible sample preparation and a new data-independent acquisition (DIA)-based MS method. Comparing with standard data dependent acquisition (DDA) approach, the optimized DIA method doubled the number of detected peptides and proteins with better reproducibility. Without protein immunodepletion and prefractionation, the single-run DIA analysis enables quantitative profiling of over 300 proteins with 50 min gradient time. The quantified proteins span more than five orders of magnitude of abundance range and contain over 50 FDA-approved disease markers. The workflow allowed us to analyze 20 serum samples per day, with about 358 protein groups per sample being identified. A proof-of-concept study on renal cell carcinoma (RCC) serum samples confirmed the feasibility of the workflow for large scale serum proteomic profiling and disease-related biomarker discovery.

Keywords:
Serum proteomic profiling
Data independent acquisition
Biomarker

1. Introduction

Blood serum or plasma is the predominant specimen for clinical proteomic studies as it can be easily accessed and contains measurable protein biomarkers that uncover physiological and pathological changes associated with diseases [1]. With recent advancement in instrumentation, mass spectrometry (MS) has been capable of providing comprehensive proteomes profiling of complex biological samples in hours. However, MS-based serum or plasma proteomics is still challenging for its high complexity and dynamic range of protein abundances. To reconcile these problems, serum samples are often extensively processed for reducing sample complexity before liquid chromatography MS (LC-MS) analysis.
These processes typically include the immunoaffinity-based depletion of high abundance proteins and extensive fractionation of protein or peptide samples [2,3]. However, immunodepletion could lead to bias because of the removal of lower abundant proteins along with the abundance species such as albumin or the antibodies used [4,5]. Furthermore, the commercial depletion kit or immunodepletion column is generally expensive which increase the detection costs. Pre-fractionation of serum sample into several or dozens of fractions by offline separation techniques such as reversed phase chromatography at basic pH and strong cation exchange chromatography (SCX) are currently widely used approach to increase the detection depth [6]. Using such strategies, it has now become possible to identify > 1000 or even > 5000 proteins in plasma [6–8]. However, the extensive pre-fractionation decreases throughput and increase pre-analytical variability, which is undesirable in clinical application. Due to the laborious sample preparation process and limited throughput, most studies in blood biomarker discovery by MS have been performed on a small number of samples, compromising the reliability of the discovered biomarkers [9]. Simplifying the pre-analysis process and developing fast, accurate, and reproducible serum proteomic profiling approach capable of analyzing large cohorts is therefore urgently needed.

Recently, Mann’s group has introduced an attractive pipeline for “plasma proteome profiling” [10,11]. Starting from 1 μL of plasma samples, the workflow enables quantitative analysis of about 300 of plasma proteins with 20 min gradients. The workflow was based on the widely used data dependent acquisition (DDA) approach for discovery and uses MaxQuant software in combination with a matching library for label-free quantitative analysis. In DDA mode, the most abundant peptide precursor ions detected in a survey scan are selected and fragmented sequentially for tandem MS identification. Although various techniques such as dynamic exclusion and charge exclusion have been developed to improve its efficiency, major drawback of the DDA mode is the loss of low abundance peptide information for accurate quantification which is largely caused by the random selection of precursor ion and non-uniform distribution of scan points on the chromatography profiles [12,13]. In contrast to DDA, recently developed data-independent acquisition (DIA) approach has opened a new dimension for quantitative proteomics [14]. The DIA mode divides the entire mass range into several sequential windows (typically 25 Da or more) and acquires all the precursors as well as fragments in the individual window without losing the low abundant ions. With this fixed cycle time acquisition feature, the DIA data sets generate a permanent digital map representing the MS-measurable proteome of the sample. Once the data was acquired, they could be easily reexamined for newly postulated biomarker validation. To date, DIA has been successfully used to study the disease-related cell lines [15], tissue [16–18], urine specimens [19], and plasma samples [20].

In this study, we introduce a streamlined workflow for serum proteomic profiling from 1 μL blood serum (Fig. 1a). We developed a new variable window DIA method specific for serum samples. Without protein depletion and pre-fractionation, the single-run DIA method allows the quantification of over 300 serum proteins with 50 min gradient time. By adopting our newly developed SISPROT technology (Simple and Integrated Spintip-based PROteomics Technology), 1 μL of blood serum sample could be processed with high sensitivity and reproducibility in a simple and fully integrated manner [21]. We successfully shortened the whole workflow to < 3.5 h from sample preparation to data analysis. As a proof of concept study, the integrated DIA-based proteomic workflow was subsequently applied to the analysis of serum samples from renal cell carcinoma (RCC) patients.

After coagulating at room temperature, the samples were centrifuged at 3000 × g for 10 min at 4 °C. Sera were stored at −80 °C prior to any further sample preparation or analysis. Ten RCC patients and 10 healthy volunteers were enrolled in this study. All the patients were diagnosed with histopathology examination and none had received chemotherapy, radiation, or undergone nephrectomy before sample collection. For spectral library construction, a pooled serum sample was prepared by pooling and mixing the same volume of each sample.

2.2. Sample preparation

Samples were prepared using the in-house developed SISPROT protocol [21] with optimization for blood serum. In brief, 1 μL of blood serum were diluted with water into 10 μL. Four microliter of diluted serum (~20 μg of protein) were acidified to pH 2–3 by 0.1% (v/v) formic acid and loaded onto a homemade SISPROT tip by centrifugation rapidly. The SISPROT was fabricated by packing 5 plus of C18 disk (3 M Empore, U.S.A.) into a standard 200 μL pipet tip and then introducing 2 μg of 20 μm POROS SCX beads (Applied Biosystems, U.S.A.). After sample loading, the proteins were washed with 20% (v/v) acetonitrile (ACN) in 8 mM potassium citrate buffer (pH 3) and then reduced by infusing 10 mM Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) in 9 mM potassium citrate buffer (pH 3) into the tip with a syringe and incubating for 15 min at room temperature. The tip was then washed with water before infusing with 2 μg/μL trypsin (Promega, Madison, WI) in 10 mM iodoacetamide, 100 mM Tris–HCl (pH 8), and incubating for 60 min at room temperature (in darkness). The digested peptides were then transferred from the SCX disk to C18 disk with 200 mM ammonium formate (pH 10). After desalting wash with 5 mM ammonium formate (pH 10), the peptides were directly eluted with 40 μL of 80% (v/v) ACN in 5 mM ammonium formate (pH 10). The eluted peptide samples were lyophilized to dryness and redissolved in 0.1% (v/v) formic acid (FA) in water. The iRT peptides (Biognoys, Schlieren, Switzerland) were spiked into the sample prior to analysis according to manufacturer instructions. All the sample preparation steps were achieved within 2 h and can be easily multiplexed on a standard centrifuge with good reproducibility.

For building a comprehensive spectral library, high-abundance protein depletion and peptide pre-fractionation were carried out for the pooled serum sample. The 12 highest abundance proteins in serum were removed by the Pierce™ TOP 12 Abundant Protein Depletion Spin Columns according to the user manual. The high pH reversed-phase pre-fractionation was performed in the SISPROT device described above, the digested peptides were fractionated by a stepwise increasing gradient of ACN (3, 6, 9, 15, and 80%) in 5 mM ammonium formate (pH 10) instead of direct elution by 80% (v/v) ACN.

2.3. DDA analysis and database search

For generation of spectral library, DDA analysis was performed on a Q Exactive mass spectrometer (Thermo Fisher Scientific) equipped with an EASY-nLC 1000 system (Thermo Fisher Scientific). The LC separation was carried out on a homemade analytical column with integrated spray tip (75 μm i.d. × 20 cm) packed with 1.9 μm/120 Å ReproSil-Pur C18 resins (Dr. Maisch GmbH, Ammerbuch, Germany) at a flow rate of 250 nL/min. The buffers used for separation were 0.1% (v/v) FA in water (buffer A) and 0.1% (v/v) FA in ACN (buffer B). Peptides were separated with a 80 min segmented gradient as follows: 3%–7% buffer B in 2 min, 7%–22% buffer B for 50 min, 22%–35% buffer B for 10 min, 35%–90% buffer B in 2 min, followed by a 16 min 90% wash. For each high pH fraction, the separation gradient time was decreased to 30 min from 7 to 22% buffer B and 6 min from 22 to 35% buffer B, respectively. The mass spectrometer was operated in data-dependent TOP10 mode. The full scan was acquired from m/z 350 to 1550 with a resolution of 70,000; MS/MS scans were performed at a resolution of 17,500 with an isolation window of 1.6 Da and higher energy collisional dissociation.
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات