JOURNAL OF PROTEOMICS 78 (2013) 389–397

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot

Analysis of human serum phosphopeptidome by a focused database searching strategy

Jun Zhu[a], Fangjun Wang[a], Kai Cheng[a], Chunxia Song[a], Hongqiang Qin[a], Lianghai Hu[b], Daniel Figeys[c,d], Mingliang Ye[a,*], Hanfa Zou[a,**]

[a]CAS Key Lab of Separation Sciences for Analytical Chemistry, National Chromatographic Research and Analysis Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China
[b]College of Life Science, Jilin University, Changchun 130023, China
[c]Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, Canada
[d]Department of Chemistry, University of Ottawa, Canada

ARTICLE INFO

Article history:
Received 8 June 2012
Accepted 8 October 2012
Available online 16 October 2012

Keywords:
Human serum
Phosphopeptidome
Ti (IV)-MCM-41
Focused database

ABSTRACT

As human serum is an important source for early diagnosis of many serious diseases, analysis of serum proteome and peptidome has been extensively performed. However, the serum phosphopeptidome was less explored probably because the effective method for database searching is lacking. Conventional database searching strategy always uses the whole proteome database, which is very time-consuming for phosphopeptidome search due to the huge searching space resulted from the high redundancy of the database and the setting of dynamic modifications during searching. In this work, a focused database searching strategy using an in-house collected human serum pro-peptidome target/decoy database (HuSPep) was established. It was found that the searching time was significantly decreased without compromising the identification sensitivity. By combining size-selective Ti (IV)-MCM-41 enrichment, RP–RP off-line separation, and complementary CID and ETD fragmentation with the new searching strategy, 143 unique endogenous phosphopeptides and 133 phosphorylation sites (109 novel sites) were identified from human serum with high reliability.

1. Introduction

The human serum peptidome, which contains thousands of endogenous peptides either degraded from proteins or secreted from cells and tissues, has been extensively investigated [1–5]. It is believed that the studies on human serum peptidome hold great promise for finding potential biomarkers of serious diseases such as cancers [6]. It should be mentioned that these endogenous peptides also have a number of post-translational modifications (PTMs), i.e., phosphorylation, glycosylation, and acetylation [5,7], which make the human serum peptidome more complex. Among these PTMs, phosphorylation is considered as one of the most important and plays a vital role in a series of biological processes including growth, metabolism and

Abbreviations: PTM, post-translational modification; CID, collision-induced dissociation; ETD, electron-transfer dissociation; MSA, multi-stage activation; FDR, false discovery rate; HuSPep, human serum pro-peptidome target/decoy database; PPD, Plasma Proteome Database; SCX, strong cation exchange; IMAC, immobilized metal ion affinity chromatography; Ti (IV)-IMAC, titanium (IV) based immobilized metal ion affinity chromatography; ACN, acetonitrile; TFA, trifluoroacetic acid; FA, formic acid; NH4FA, ammonium formate; AP, alkaline phosphatase; M.W., molecular weight.

* *Corresponding author*. Tel./fax: +86 411 84379620.
** *Corresponding author*. Tel.: +86 411 84379610; fax: +86 411 84379620.
E-mail addresses: mingliang@dicp.ac.cn (M. Ye), hanfazou@dicp.ac.cn (H. Zou).

1874-3919/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jprot.2012.10.006

cell differentiation [8]. Therefore, serum phosphopeptidome, as an intersection of peptidome and phosphoproteome, might be an important source for retrieving diagnostic information of diseases.

Recently, human serum phosphoproteome has already been explored by several labs. Zhou et al. enriched phosphopeptides from tryptic digest of serum sample by TiO_2 and identified about 100 unique phosphopeptides after LC–MS/MS analysis [9]. Meanwhile, Carrascal et al. applied strong cation exchange (SCX) for fractionation of phosphopeptides enriched from serum protein digest. After LC–MS/MS analysis, multiple search engines were applied to search the MS data. They identified 138 unique phosphopeptides, among which, 85 were identified by more than one search engine [10]. Garbis et al. combined size exclusion fractionation, hydrophilic interaction chromatography and ultraperformance RPLC–MS/MS to analyze the serum proteome and phosphoproteome of patients with benign prostate hyperplasia [11]. They identified 375 phosphopeptides at FDR < 5%. These studies indicated the presence of phosphorylated proteins in serum though the number is much smaller than that in tissues. In addition to these phosphoproteomic studies, the detection of endogenous phosphopeptides in human serum was also reported. Cirulli et al. analyzed endogenous phosphopeptides in different biological fluids including urine, saliva and serum. They only identified 4 endogenous phosphopeptides derived from fibrinogen in serum [12]. Zhang et al. applied cerium ion-chelated magnetic silica microspheres for enrichment of human serum endogenous phosphopeptides and only 3 from fibrinogen were detected [13]. In our previous study, Ti (IV)-MCM-41 was synthesized and successfully used for direct enrichment of endogenous phosphopeptides from human serum, and also only 4 phosphopeptides were detected after MALDI analysis [14]. In spite of these studies, the number of endogenous phosphopeptides identified in human serum is still limited.

A main challenge for phosphopeptidome analysis is that current database searching strategy using proteome database is not efficient and sensitive for endogenous phosphopeptide identification. Endogenous peptides detected in peptidomic studies are often derived from the combined activity of different kinds of peptidases, some of which are of unknown specificity. When searching for endogenous peptides using a search engine like Mascot or Sequest, no enzymatic restriction is set. This makes the pool of peptide candidates for a particular MS/MS spectrum very large when a proteome database is adopted. It is even larger for identification of endogenous phosphopeptides because there are three more variable phosphorylated modifications on Ser, Thr and Tyr. Due to the huge search space, such a search strategy is extremely slow and insensitive. A straight forward approach to solve this problem is to reduce the search space by using small databases with low redundancy.

Several peptidome databases have been developed recently, such as PepBank [15] and SwePep [16,17]. PepBank is a database of peptide sequences and SwePep is a database of large collections of peptide spectra, both holding great potential in peptidome studies of many biological samples such as brain. In serum, majority of endogenous peptides are derived from protein degradation due to the coexistence of a variety of active peptidases. The serum proteins could be cut by endopeptidases and the resulting peptides could be further digested by exopeptidases to produce a series of peptide ladders. The databases mentioned above do not include all of these peptides and are not suited for serum peptidome study.

In this work, a small focused database, named as human serum pro-peptidome database (HuSPep), was specially generated for human serum phosphopeptidome analysis. It was found that over 95% of endogenous peptides which were identified by searching against IPI human database could be obtained by using HuSPep. Then, HuSPep was applied for in-depth analysis of human serum phosphopeptidome. Human serum endogenous phosphopeptides were directly enriched by size-selective Ti (IV)-MCM-41 material, and analyzed by off-line RP–RP separation and MS/MS detection with MSA (multi-stage activation, CID) and ETD. The acquired spectra were then searched against HuSPep for initial endogenous phosphopeptide identification. The spectra for these initial identified phosphopeptides were further searched against IPI human target/decoy database. The final results were achieved by keeping the consistent identifications from both searching. By this way, 143 unique endogenous phosphopeptides with 133 phosphorylation sites in human serum were identified with high confidence.

2. Materials and methods

2.1. *Materials*

Protease inhibitor cocktail, alkaline phosphatase, trifluoroacetic acid (TFA), and (3-aminopropyl) trimethoxysilane were purchased from Sigma (St. Louis, MO). PhoSTOP (phosphatase inhibitor cocktail tablets) was purchased from Roche (Mannheim, Germany). Formic acid (FA) was obtained from Fluka (Buches, Germany). Acetonitrile (ACN, HPLC grade) and 25% ammonia solution were purchased from Merck (Darmstadt, Germany). Pure water used in all the experiments was purified with a Milli-Q system (Millipore, Milford, MA). Other chemicals used were either of analytical grade or better.

2.2. *Sample preparation*

The human serum used in all the experiments of this study was obtained from the Second Affiliated Hospital of Dalian Medical University (Dalian, China) and was a pool of 100 serum samples from healthy individuals. The utilization of human serum complied with guideline of Ethics Committee of the Hospital. The human serum was stored at −80 °C until further use.

Ti (IV)-MCM-41 was prepared as reported previously [14]. Serum endogenous phosphopeptides were enriched with Ti (IV)-MCM-41 and the enrichment procedure was similar to that in our previous reports [14,18]. Briefly, 500 μL of human serum was diluted with 30% ACN/0.1% TFA to a final volume of 5 mL, protease inhibitor and phosphatase inhibitor were both added to 0.5% (v/v). Then, 5 mg of Ti (IV)-MCM-41 was added to the serum, followed by incubation at 4 °C using a rolling incubator (Co., Ltd, Haimen) for 40 min. After centrifugation at 20,000 *g* for 4 min, the supernatant was collected and enriched with 2 mg Ti (IV)-MCM-41 for a second time to ensure that all the phosphopeptides were captured. The Ti