



Label-free relative quantification method for low-abundance glycoproteins in human serum by micrOTOF-Q

Piliang Hao^{a,*}, Yan Ren^a, Yongming Xie^b

^a Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing Airport Industrial Zone B-6, Beijing 101318, China

^b Bruker Daltonics Inc. (China), Beijing, China

ARTICLE INFO

Article history:

Received 19 November 2008

Accepted 4 April 2009

Available online 14 April 2009

Keywords:

Relative quantification

Glycoprotein

Serum

Depletion

Immunoaffinity

Multi-lectin column chromatography

micrOTOF-Q

Mass spectrometry

Proteomics

ABSTRACT

In this study, a label-free relative quantification strategy was developed for quantifying low-abundance glycoproteins in human serum. It included three steps: (1) immunodepletion of 12 high-abundance proteins, (2) enrichment of low-abundance glycoproteins by multi-lectin column, (3) relative quantification of them between different samples by micrOTOF-Q. We also evaluated the specificity and efficiency of immunodepletion, the accuracy of protein quantification and the possible influence of immunodepletion, glycoprotein enrichment, trypsin digestion and peptide ionization on quantification. In conclusion, the relative quantification method can be effectively applied to the screening of low-abundance biomarkers.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Glycosylation is one of the most common post-translational modifications (PTMs), which plays an important role in many biological processes, such as embryonic development, cell-to-cell interactions, immunological reactions and cell division processes [1]. It is reported that both the expression level and glycosylation status of some glycoproteins change significantly in the sera, tissues or cerebrospinal fluids of human with severe diseases, such as cancer, sepsis, inflammation and so on [2–5]. And some of them have been used as biomarkers in the early diagnosis of cancers with high sensitivity and specificity clinically, such as alpha-fetoprotein (AFP), CA-125 and prostate-specific antigen (PSA) [6].

Compared with cerebrospinal fluid or tissue samples, serum is an ideal sample of seeking biomarkers for early diagnosis and the monitoring of subsequent therapy for cancer or other diseases because it is much easier to collect than others [7]. However, because the proteins in human serum are of a very broad dynamic range, the analysis of low-abundance proteins in serum is generally masked by high-abundance proteins, which constitute over 90% of the total protein content of the serum [8]. To solve this problem, many immunoaffinity depletion kits were developed, such as Mul-

tipule Affinity Removal System (MARS) from Agilent [9], Seppro[®] IgY12 Liquid Chromatography Column from Genway Biotech and ProteoPrep[®] 20 Immunoaffinity Depletion Kit from Sigma. In the past several years, with the help of the above-mentioned depletion kits, many efforts were made to look for biomarker candidates in cancer serum samples using 2-DE based proteomic methods [10–13], but most of the differential proteins found were relatively high-abundance proteins that could not be used as biomarkers clinically because it was rather difficult for 2-DE to analyze low-abundance proteins. However, only small amounts of differentially expressed proteins were released into blood in the early stage of diseases. As we knew, most of the currently usable biomarkers were relatively low-abundance glycoproteins. Thus, the relative quantification of low-abundance glycoproteins between test and control samples would be extremely helpful to the early diagnosis of diseases. Yang and Hancock developed an alternative method to isolate glycoproteins from high-abundance proteins, i.e. multi-lectin affinity chromatography (M-LAC) method, which could enrich and identify N-linked and O-linked glycoproteins [14,15]. It indicated that M-LAC was an efficient tool for enriching glycoproteins and depleting albumin.

In recent years, stable isotope labeling methods were widely used in relative quantification because of their high accuracy and sensitivity, but they were less suitable for serum samples because increased side reactions occurred in the chemical derivatization of low-abundance peptides [16]. And the cost of labeling methods is

* Corresponding author.

E-mail addresses: hpiliang@163.com, haopl@big.ac.cn (P. Hao).

also much more expensive than that of label-free quantification methods. Thus, Yang and Hancock also used M-LAC together with LTQ-FTMS in the label-free relative quantification of differentially expressed glycoproteins in human breast cancer serum samples [17]. And they used the extracted ion chromatogram (EIC) of the m/z signal of the peptide with a mass window of 0.5 Da for label-free relative quantification. Because of the limitation of the mass accuracy, it would affect the accuracy of relative quantification somewhat. And the remaining high-abundance glycoproteins except for albumin would also affect the analysis of low-abundance proteins.

In this study, we added a given amount of horseradish peroxidase (HRP) as the internal standard in serum at the beginning. And then we depleted the top 12 proteins from human serum with Seppro[®] IgY12 Liquid Chromatography Column, enriched the low-abundance glycoproteins with M-LAC (ConA/WGA) from the flow-through of IgY12 column and used micrOTOF-Q from Bruker Daltonics to analyze them after trypsin digestion. Finally, because of the high mass accuracy of micrOTOF-Q, high resolution extracted ion chromatograms (hrEICs) with a mass window of 0.02 Da were generated, and their peak area values were used for quantification. We also evaluated the repeatability, specificity and efficiency of the immunodepletion, the accuracy of quantification by micrOTOF-Q and the possible influence of immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization on protein relative quantification. Based on our results, the relative quantification method presented in this paper can be effectively applied to the screening of low-abundance glycoprotein biomarker candidates in human serum.

2. Experimental

2.1. Materials

The prepacked Seppro[®] IgY12 Liquid Chromatography Column (MIXED-LC5) was purchased from GenWay Biotech (San Diego, CA). HPLC-grade acetonitrile was purchased from Fisher Scientific (Fairlawn, NJ). HPLC-grade formic acid was purchased from Dima Technology Inc. Agarose bound Wheat germ agglutinin (WGA, L1394-5ML), agarose bound Concanavalin A (Con A, C7555-5ML), Methyl α -D-mannopyranoside (M6882-25G), *N*-acetyl-D-glucosamine (A8625-5G) and Protease Inhibitor Cocktail Set I (P8340) were purchased from Sigma–Aldrich. Modified trypsin (sequence grade) was purchased from Promega (Madison, WI, USA). Microcon YM-10 (0.5 ml, 42406), Amicon Ultra-15 10 kDa Centrifugal Filters (15 ml, UFC901008) and Millex[®] GP 0.22 μ m filters were purchased from Millipore. Albumin/IgG Removal Kit (122642) was purchased from Calbiochem. Horseradish Peroxidase (HRP, P121606) was purchased from Roche. C18 packing (Synergi, 04G-4375) was from Phenomenex. Human sera were taken from 10 healthy donors. All other materials were purchased from Sigma unless otherwise described.

2.2. Immunodepletion of 12 high-abundance proteins

Forty microliters of human serum was collected from each of 10 human sera and mixed thoroughly. Two hundred microliters of the pooled sera were centrifuged at 12,000 g for 10 min and diluted 15 times with 2800 μ l dilution buffer (10 mM Tris, pH 7.4, 0.15 M NaCl). And then 5 μ l 4 mg/ml HRP and 15 μ l Protease Inhibitor Cocktail were added into the diluted serum. Subsequently, it was forced through a 0.22 μ m filter by a 5 ml syringe. The Seppro[®] IgY12 column enables the removal of albumin, IgG, α 1-antitrypsin, IgA, IgM, transferrin, haptoglobin, α 1-acid glycoprotein, α 2-macroglobin, HDL (apolipoproteins A-I and A-II) and fibrinogen in one step [8]. The immunodepletion procedure followed the protocol pro-

vided by Genway Biotech. About 900 μ l diluted serum was loaded once. Stripping buffer (0.1 M glycine, pH 2.5) and neutralizing buffer (0.1 M Tris, pH 8.0) were also mandatory for the depletion procedure. After the immunodepletion, the final volume of the removed high-abundance proteins in stripping buffer and the low-abundance proteins in dilution buffer was about 30 ml and 18 ml, respectively. The high-abundance proteins were solvent exchanged with 50 mM ammonium bicarbonate buffer in a 10 kDa Microcon YM-10 Filter (0.5 ml). The low-abundance proteins were concentrated to a final volume of about 500 μ l with Amicon Ultra-15 10 kDa Centrifugal Filters (15 ml) before use. The amount of loaded proteins and the collected fractions was determined by Bradford assay.

2.3. Preparation of multi-lectin affinity chromatography (M-LAC) columns

The M-LAC was prepared by mixing 100 μ l of agarose bound Con A and 100 μ l of agarose bound WGA in a dismantled Albumin/IgG Removal Column (122642, Calbiochem). The column was then agitated slightly to mix the gel thoroughly. No bubbles should stay in the gel because they would affect the enrichment efficiency of glycoproteins. The flow through the column was gravity driven. It could not be reused in order to avoid the crossover contamination between different samples. Before samples were loaded, the column was washed with 1–2 ml wash solution (0.15 M NaCl, 20 mM Tris, 5 mM CaCl₂, 5 mM MgCl₂, pH 7.4).

2.4. Enriching low-abundance glycoproteins with M-LAC

The prepared M-LAC column was equilibrated with 1–2 ml equilibration buffer (0.5 M NaCl, 20 mM Tris, pH 7.4). And then 500 μ l of the concentrated low-abundance proteins obtained in Section 2.2 were loaded into the column. After 20 min incubation at room temperature, the unbound proteins were washed with 3–5 ml equilibration buffer until no proteins were eluted from the column. And then the bound proteins were eluted with 1.2 ml elution buffer (0.5 M NaCl, 20 mM Tris, 0.36 M Methyl α -D-mannopyranoside, 0.25 M *N*-acetyl-glucosamine, pH 7.4). The eluted fractions were concentrated into 20 to 50 μ l with Microcon YM-10 Centrifugal Filters (0.5 ml). Subsequently, 500 μ l 50 mM ammonium bicarbonate buffer was added into the filters and concentrated into the final volume of 20–50 μ l again. And then another 500 μ l 50 mM ammonium bicarbonate buffer was added into the filters and concentrated into the final volume of about 30 μ l. The amount of loaded proteins and the collected fractions was determined by Bradford assay.

2.5. Trypsin digestion

Both the solution of high-abundance proteins and low-abundance glycoproteins obtained in Section 2.2 and Section 2.4 were dried in vacuum. And then the proteins were redissolved in 6 M urea in water, reduced with 5 mM DTT at 56 °C for 1 h and alkylated with 20 mM iodoacetamide for 45 min in the dark. The concentration of urea was diluted to less than 1 M with 50 mM ammonium bicarbonate buffer, and trypsin was added into the solution at a rate of 1:60 (trypsin: samples). And then it was incubated at 37 °C for 8–12 h. For complete digestion, it continued to be incubated for about 12 h at room temperature after the same amount of trypsin was added.

2.6. LC-MS/MS

The trypsin digested peptides were separated and analyzed on Ultimate 3000 nanoHPLC (DIONEX) coupled to micrOTOF-Q (Bruker Daltonics). Approximately 2 μ g of each sample were

injected into the trap column with the autosampler of Ultimate 3000. And then the trap column was washed with 0.1% formic acid at a flow rate of 20 $\mu\text{l}/\text{min}$ for 5 min in order to desalt the samples. Subsequently, the 10-port valve was switched to direct the flow to the separation column. The desalted peptides were then separated on a C18 column (packed in-house; Synergi C18; 150 mm \times 0.075 mm) and analyzed on micrOTOF-Q mass spectrometer with a nanoelectrospray ionization ion source. The flow rate was maintained at 400 nl/min. The gradient was started at 3% acetonitrile (ACN) with 0.1% formic acid and linearly increased to 43% ACN in 80 min, then to 73% ACN in 12 min, and to 95% ACN in another 13 min. The gradient was then decreased to 3% ACN in 1 min and maintained at 3% ACN for 14 min. Each sample was subjected to two separate MS analysis: one tandem MS (MS/MS) for identification of proteins and one full mass scan for quantitative analysis. Both the MS analysis was repeated once. The temperature of the ion transfer tube was set at 165 $^{\circ}\text{C}$, and the spray voltage was 1.4 kV. The collision energy of the quadrupole was set at 35 eV/z for MS/MS. Data-dependent ion selection was monitored to select the most abundant two ions from an MS scan for MS/MS analysis. Dynamic exclusion was continued for 2 min.

2.7. Bioinformatics

Mass spectra were processed with DataAnalysis 3.4 from Bruker Daltonics, and the resulting MGF documents were searched for tryptic peptides with up to one miscleavage against SwissProt human database using Mascot software (Matrix Science Ltd.). The fixed modification was set as carbamidomethylation (C), and variable modifications were set as oxidation (M) and pyro-Glu (N-term Q). Both the peptide and MS/MS mass tolerance was set as ± 0.04 Da.

2.8. Relative quantification by the peak areas of hrEICs

The hrEICs of the identified peptides were generated with a mass window of 0.02 Da. Because only two of the most abundant ions were selected for MS/MS fragmentation, the identification of some peptides might be omitted. Thus, if the hrEIC peak of an identified peptide was found in the full mass scan within 0.5 min of the same retention time when the peptide was identified, it would be used for peak area measurement for the peptide no matter whether the peptide was identified in the sample or not.

To normalize the possible influence of immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization on protein quantification, same amount of HRP was added into all samples at the beginning as an internal standard. Five identified peptides of HRP with highest scores were used for peak area measurement. The individual peak area ratios of each peptide between different samples were calculated separately, and the average of them was used as the ratio of HRP, and the relative standard deviation was also obtained. If one of the calculated ratios was over 50% away from the average of others, it would be excluded from the calculation of averages. It was statistically reasonable to exclude the values far away from others because some unexpected factors might affect the ionization efficiency of certain peptides and lead to the different calculated ratios although the LC-MS/MS runs were highly reproducible. If a protein was identified with only two peptides, but with quite different calculated ratios, a third run was necessary to determine the average for the two peptides. If the calculated ratios of them were highly reproducible and still far away from each other in the three runs, the calculated ratios should be accurate, and then other validation methods should be employed to verify the ratio of the protein, such as western blot. The relative quantification of the identified proteins between different samples was similar to that of HRP. If a protein was identified with over three peptides, the three identified peptides with highest scores would be used for

relative quantification. If not, all of the identified peptides would be used. Based on our calculation, it was of no significant difference even only three identified peptides with high scores were used for relative quantification, but it significantly reduced the workload because certain proteins were identified with over 30 peptides. The final ratio of a protein between different samples was equal to the calculated ratio of the protein divided by that of HRP.

2.9. Evaluation of the detection limit, accuracy and linear range of protein quantification

We loaded 10 fmol Bovine Serum Albumin (BSA) as the standard before the MS analysis on samples was performed. If over four unique peptides were identified with a positive score, we continued to analyze samples. To evaluate the accuracy of the label-free quantification, 2 μl and 3 μl low-abundance glycoproteins were loaded respectively, and both the MS and MS/MS analyses were conducted in triplicate. And then we quantified all of the identified proteins using the hrEICs of their respective identified peptides. If the relative quantification of all proteins in the low-abundance fractions was very accurate, the concentration of them would be within the linear range of quantification.

2.10. Evaluation of the possible influence of immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization on protein quantification

To evaluate the possible influence of immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization on protein quantification, we added about 0.2% and 0.4% HRP into two aliquots of the same serum sample at the beginning. And then they were subjected to the same treatment in our strategy. Samples with different concentration of HRP were prepared in duplicate, and both the MS and MS/MS analyses were also conducted in duplicate. Thus, if the relative ratio of HRP between them was about 2, we believed that the above-mentioned factors had no significant influence on the relative quantification of HRP. After normalization with HRP, if the relative ratio of other proteins between the two aliquots was about 2, the above-mentioned factors also had no significant influence on the relative quantification of them.

3. Results and discussion

The strategy used in this work was illustrated in Fig. 1. To normalize the possible influence of immunodepletion, enrichment of

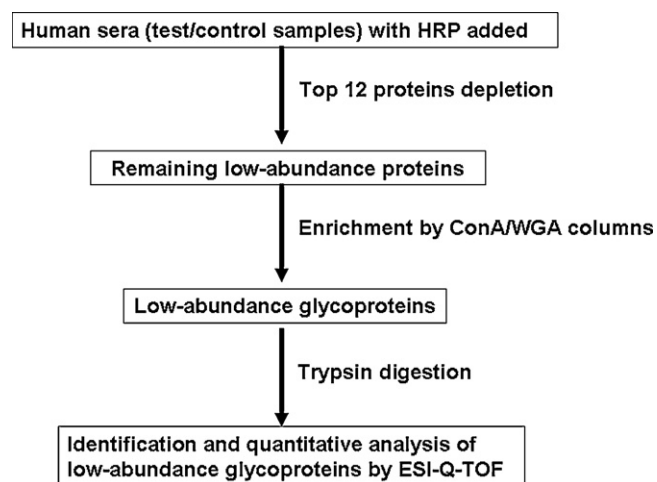


Fig. 1. The strategy for label-free relative quantification of low-abundance glycoproteins in human serum by micrOTOF-Q. For details, see Section 2.

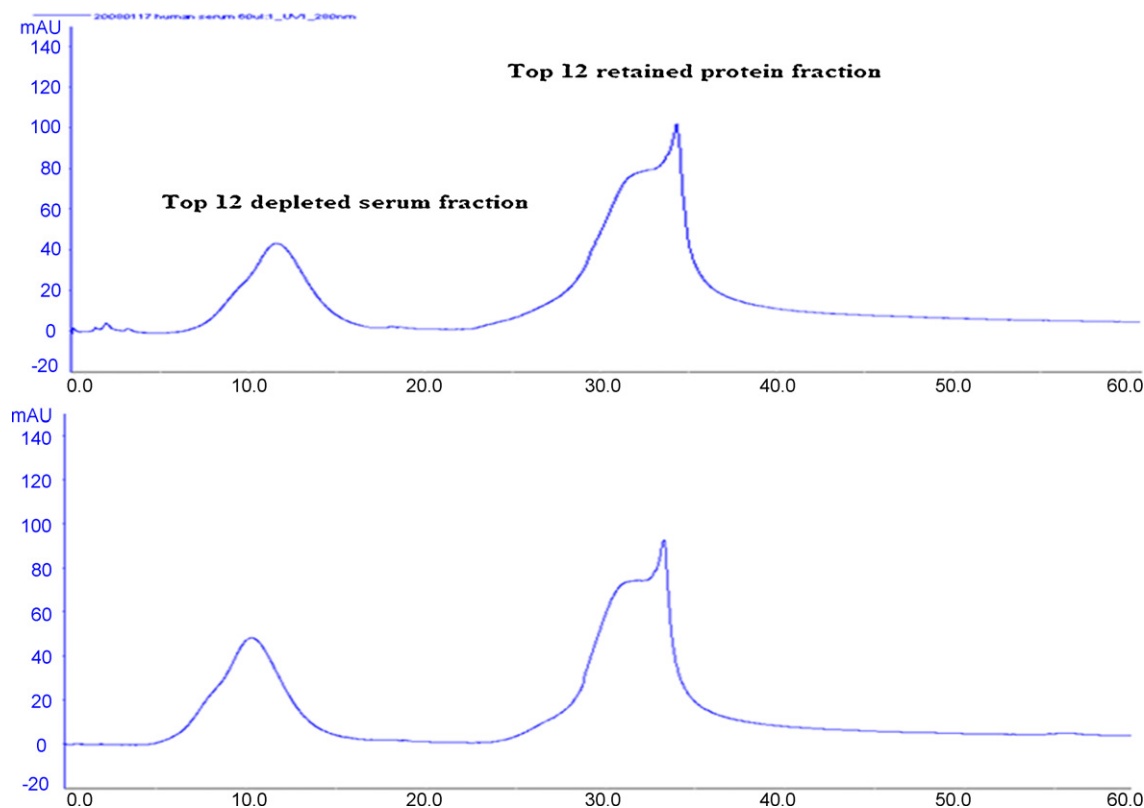


Fig. 2. UV Chromatograms of two sequential depletions of 60 μ l of the same serum. During the immunodepletion process, the absorption was set at 280 nm. The flow-through of the column was collected as the top 12 depleted fraction, and the eluted fraction was collected as the top 12 proteins. The two chromatograms were highly reproducible.

glycoproteins, trypsin digestion and peptide ionization on protein quantification, same amount of HRP (about 0.2% of the total protein content of the sample) was added into all samples at the beginning as an internal standard. And then top 12 proteins were depleted with Seppro[®] IgY12 columns in order to improve the detection of low-abundance proteins. The repeatability, specificity and efficiency of the immunodepletion were evaluated. The flow-through of the column was then concentrated and used for the enrichment of low-abundance glycoproteins with ConA/WGA columns. After digested by trypsin, the enriched glycoproteins were subjected to MS/MS analysis and full mass scan twice. The hrEICs of the identified peptides of each protein with a mass window of 0.02 Da were generated, and their peak area values were used for quantification. We also evaluated many factors that might interfere with the relative quantification in this paper, such as immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization, and concluded that they had no significant influence on it.

3.1. Analysis of the repeatability, specificity and efficiency of Seppro[®] IgY12 immunodepletion

Top 12 high-abundance proteins represented about 93% of total protein, which greatly masked the identification and quantification of other low-abundance proteins [18,19]. Because only small amounts of differentially expressed proteins were released into blood in the early stage of diseases, the depletion of high-abundance proteins was obligatory for the monitoring of their changes. Many kinds of immunodepletion columns had been utilized during the analysis of serum samples. Although the performance of different kinds of immunodepletion columns was evaluated comprehensively in the past [8,9,20], we still carefully evaluated the repeatability, specificity and efficiency of Seppro[®] IgY12 in this

work because our aim was to quantify low-abundance glycoproteins between normal and diseased samples, and it was crucial to ensure that the immunodepletion would not lead to the change of protein concentration.

As shown in Fig. 2, the UV chromatograms of two sequential depletions of 60 μ l of the same serum were highly reproducible. The concentration of each obtained fraction was also determined by Bradford assay. For the immunodepletion, after 4.30 mg diluted human serum was loaded, 3.79 mg of high-abundance proteins and 285 μ g of low-abundance proteins were collected. The low-abundance proteins were then loaded onto M-LAC columns after concentration. Subsequently, 65 μ g glycoproteins and 170 μ g flow-through were collected. The recovery of immunodepletion and glycoprotein enrichment was about 95% and 82% respectively, and they were consistent among different runs. In order to prove the specificity of the immunodepletion, the high-abundance fractions were also analyzed with microTOF. Table 1 showed that most of removed high-abundance proteins were the indicated top 12 proteins from Genway, and only a few of them were out of them. Even the commonly known relatively higher-abundance proteins, such as ceruloplasmin and complement C3, did not appear in the removed fraction. It indicated that the immunodepletion procedure was highly specific. However, Rapkiewicz AV et al reported that the immunodepletion procedure could lead to the loss of potential biomarkers [21]. The possible explanation that no potential biomarkers were identified in the removed fraction in our experiment was due to their extremely low concentration and the dynamic range of detection. Although some low-abundance proteins might be lost during the immunodepletion, it remained one of most efficient ways to study low-abundance proteins in serum samples because no alternative methods could provide better overall performance until now. Bradford assay result indicated that the flow-through of the immunodepletion column owned about 7% in

Table 1
Identification list of the removed top 12 proteins by micrOTOF-Q.

No.	Protein description ^a	Accession no. ^b	Matched peptides	Sequence coverage	Protein score	Glycoprotein ^c
1	Serum albumin precursor	spiP02768	58	66%	2419	No
2	Serotransferrin precursor	spiP02787	25	38%	1329	Yes
3	Cytokeratin-1	spiP04264	19	27%	832	No
4	Cytokeratin-10	spiP13645	13	19%	721	No
5	Apolipoprotein A-I precursor	spiP02647	16	51%	633	Yes
6	Cytokeratin-2e	spiP35908	12	19%	625	No
7	Alpha-1-antitrypsin precursor	spiP01009	13	31%	648	Yes
8	Ig gamma-1 chain C region	spiP01857	12	28%	574	Yes
9	Alpha-2-macroglobulin precursor	spiP01023.1	19	15%	881	Yes
10	Haptoglobin precursor	spiP00738	15	33%	638	Yes
11	Cytokeratin-9	spiP35527	10	13%	520	No
12	Ig mu chain C region	spiP01871	7	18%	374	Yes
13	Ig gamma-2 chain C region	spiP01859	9	23%	345	Yes
14	Hemoglobin subunit beta	spiQ6WN21	6	48%	273	No
15	Ig kappa chain C region	spiP01834	3	48%	250	No
16	Ig alpha-1 chain C region	spiP01876	4	14%	209	Yes
17	Alpha-1-acid glycoprotein 1	spiP02763	5	27%	186	Yes
18	Ig lambda chain C regions	spiP01842	3	28%	152	No
19	Ig heavy chain V region 5A	spiP19181	2	18%	114	No
20	Ig heavy chain V-III region VH26 precursor	spiP01764	2	18%	109	No
21	Transthyretin precursor	spiP02766	2	17%	102	Yes
22	Cytokeratin-14	spiP02533	2	4%	98	No
23	Ig mu heavy chain disease protein (BOT)	spiP04220	2	5%	88	No
24	Apolipoprotein A-II precursor (Apo-AII)	spiQ8MIQ5	2	10%	82	No
25	Isocitrate dehydrogenase subunit alpha	spiQ5R678	2	2%	49	No
26	Ig kappa chain V-II region Cum	spiP01614	1	11%	92	No
27	Ig kappa chain V-I region Lay	spiP01605	1	8%	53	No
28	Ig heavy chain V-I region HG3 precursor	spiP01743	1	9%	45	No
29	Ig heavy chain V-III region GAL	spiP01781	1	7%	35	No
30	Ig kappa chain V-III region VG precursor	spiP04433	1	7%	34	No
31	Succinyl-CoA ligase	spiO82662	1	2%	34	No
32	Apolipoprotein C-I precursor	spiP02654	1	10%	34	No

^a Protein description in Swiss-Prot database.

^b Accession number was recorded as a reference for the identification in Swiss-Prot database.

^c It was based on the annotation in Swiss-Prot database.

total protein, which was consistent with the results from others. In Table 2, none of the removed top 12 proteins appeared in the identification list of low-abundance glycoproteins so that they would not affect the identification and quantification of them any more. And it also indicated that the efficiency of the immunodepletion was sufficient. The identification results of both the high-abundance and low-abundance fractions from many different immunodepletion runs also indicated that the procedure was highly reproducible (data not shown). Thus, it would not bring any significant change to the relative concentration of low-abundance proteins between different serum samples, and it could be safely used in the relative quantification experiment.

Another newly developed ProteoPrep® 20 Immunoaffinity Depletion Kit from Sigma has also already been widely used in the analysis of human serum samples, and it improves the detection of low-abundance proteins by mass spectrometry. In their technical article, 51 relatively low-abundance proteins were identified in LC-MS/MS analysis, in which 11 of them were identified with only one unique peptide. In our experiment, as shown in Table 2, 62 relatively low-abundance proteins were identified, in which 17 of them were identified with only one unique peptide. In order to confirm that the identified low-abundance glycoproteins in our experiments were really low abundant in human serum, we compared them with that from the technical article of ProteoPrep 20. As expected, 36 proteins (highlighted in italics in Table 2) appeared in both of our and their results. It indicated that our strategy was also powerful in analyzing low-abundance proteins in human serum. Furthermore, because our strategy focused mainly on N-linked glycoproteins, it was especially useful in analyzing unglycosylated proteins that would have interactions with glycoproteins under diseased conditions although unspecific enrichment happened occasionally.

3.2. Evaluation of the detection limit, accuracy and linear range of protein quantification by micrOTOF

As proteomics developed quickly, many scientists began to pay much attention to quantitative analysis and elucidation of the post-translational modifications (PTMs) of biological samples [22–23]. Now, MS-based quantitative proteomics can be achieved through both label-free and stable isotope labeling methods, which are better than other quantitative methods, such as 2-DE and protein arrays, in many aspects [24]. LTQ, LTQ-FTMS and LTQ-Orbitrap have been widely used in the quantitative analysis of biological samples [16,25–26]. In this work, we utilized micrOTOF with extremely high resolving power and mass accuracy in the identification and quantification of low-abundance proteins with specific PTM, i.e. glycosylation. In 2007, Bruker Daltonics used micrOTOF in the quantification of four spiked proteins in two cell lysate samples. However, to our knowledge, it has not been used in the analysis of real complex samples, such as human serum, until now. Firstly, we should demonstrate the performance of micrOTOF in the label-free relative quantification of serum samples.

Before any MS analysis was performed, 10 fmol BSA was loaded as the standard to test the performance of micrOTOF. If over four unique peptides were identified with a positive score, we continued to analyze samples. It could not be regarded as the detection limit of micrOTOF in analysis of complex samples, but it might be around this level if the identification of a protein was validated with only one or two unique peptides identified in complex samples. In order to quantify proteins between different samples, as shown in Fig. 3, the hrEICs of identified peptides of complement C3 with a mass window of 0.02 Da were generated. Because of the high mass accuracy, it allowed for a distinguished selectivity from the background.

Table 2
Identification list of the low-abundance glycoproteins by microTOF-Q.

No.	Protein description ^a	Accession no. ^b	Matched peptides	Sequence coverage	Protein score	Glyco-protein ^c
1	Complement C3 precursor	splP01024.2	60	40%	3672	Yes
2	Cytokeratin-10	splP13645	24	42%	1784	No
3	Ceruloplasmin precursor	splP00450	24	29%	1457	Yes
4	Alpha-1-antichymotrypsin precursor	splP01011	21	48%	1271	Yes
5	Cytokeratin-1	splP04264	20	32%	1235	No
6	ITI heavy chain H4	splQ14624	17	22%	1058	Yes
7	Complement C4-A precursor	splP0C0L4.1	18	14%	1034	Yes
8	ITI heavy chain H1	splP19827	12	19%	957	Yes
9	Hemopexin precursor	splP02790	22	40%	908	Yes
10	Antithrombin-III precursor (ATIII)	splP01008	13	44%	903	Yes
11	Complement factor B precursor	splP00751	17	25%	879	Yes
12	Complement C5 precursor	splP01031.4	15	14%	840	Yes
13	Angiotensinogen precursor	splQ9GLN8	10	29%	797	Yes
14	ITI heavy chain H2	splP19823	15	20%	792	Yes
15	Alpha-1B-glycoprotein precursor	splP04217	10	32%	679	Yes
16	Plasma protease C1 inhibitor precursor	splP05155	10	24%	628	Yes
17	Heparin cofactor 2 precursor	splP05546	10	25%	634	Yes
18	Cytokeratin-9	splP35527	10	17%	616	No
19	Cytokeratin-2e	splP35908	8	17%	534	No
20	Cytokeratin-17	splQ04695	11	23%	530	No
21	Cytokeratin-6B	splP04259.4	9	17%	512	No
22	Kininogen-1 precursor	splP01042	5	11%	358	Yes
23	Complement factor H precursor	splP08603.4	7	6%	321	Yes
24	Afamin precursor	splP43652	5	14%	284	Yes
25	Alpha-2-HS-glycoprotein precursor (Fetuin-A)	splQ9N2D0	5	25%	267	Yes
26	Lumican precursor	splP51884	4	13%	217	Yes
27	Complement factor I precursor	splP05156	5	8%	215	Yes
28	Beta-2-glycoprotein 1 precursor	splQ95LB0	4	15%	214	Potential
29	Insulin-like growth factor-binding protein complex acid labile chain precursor (ALS)	splP35858	3	10%	184	Yes
30	Complement component C9 precursor	splP02748	3	6%	161	Yes
31	Kallistatin precursor (Serp1 A4)(Kallikrein inhibitor)	splP29622	4	11%	145	Yes
32	Alpha-2-antiplasmin precursor	splP08697	4	16%	118	Yes
33	AMBIP protein precursor	splP02760.1	2	5%	116	Yes
34	ITI heavy chain H3	splQ06033.2	3	5%	114	Yes
35	Carboxypeptidase N catalytic chain precursor (CPN)	splQ2KJ83	2	10%	112	Yes
36	Complement C2 precursor	splQ8SQ74	2	3%	111	Yes
37	N-acetylmuramoyl-L-alanine amidase precursor (PGRP-L)	splQ96PD5	2	6%	108	Yes
38	Dermcidin precursor	splP81605	2	22%	102	No
39	Apolipoprotein B-100 precursor	splP04114	5	1%	92	Yes
40	Complement component C6 precursor	splP13671	2	5%	85	Yes
41	Thyroxine-binding globulin precursor	splP05543	2	4%	77	Yes
42	Pigment epithelium-derived factor precursor (PEDF)	splP36955	2	6%	74	Yes
43	Complement C1s subcomponent precursor (C1 esterase)	splP09871	2	3%	60	Yes
44	Leucine-rich alpha-2-glycoprotein precursor (LRG)	splP02750	2	11%	48	Yes
45	Alpha-S1-casein precursor	splP02662	2	8%	37	No
46	Complement component C8 gamma chain precursor	splP07360	1	7%	93	No
47	Complement component C8 alpha chain precursor	splP07357	1	2%	81	Yes
48	Sex hormone-binding globulin precursor (SHBG)	splP04278	1	5%	72	Yes
49	Carboxypeptidase B2 precursor	splQ96IY4	1	5%	69	Yes
50	Protein S100-A9 (S100 calcium-binding protein A9)	splP06702	1	11%	66	No
51	Complement C1r subcomponent precursor	splQ5R1W3	1	2%	65	Potential
52	Prothrombin precursor	splQ5R537	1	3%	64	Yes
53	Histidine-rich glycoprotein precursor	splP04196	1	2%	60	Yes
54	Vitronectin precursor	splP04004.1	1	3%	59	Yes
55	Clusterin precursor	splP10909.1	1	2%	54	Yes

56	Protein S100-A11 (S100 calcium-binding protein A11)		1	16%	51	Potential
57	<i>Fibronectin</i>	splQ6B345	1	0%	49	Yes
58	Gamma-glutamyl hydrolase precursor (Gamma-Glu-X carboxypeptidase)	splP02751.3	1	4%	44	Potential
59	Complement component C8 beta chain precursor	splQ92820	1	2%	43	Yes
60	Complement component C7 precursor	splP07358	1	1%	40	Yes
61	Putative cytochrome c oxidase subunit II PS17	splP10643	1	50%	37	No
62	GTP-binding protein TypA/BipA homolog	splP84733.1 splP72749	1	2%	36	No

Note: The proteins of which the protein description was in italics were also identified in the technical article of ProteoPrep® 20 Immunoaffinity Depletion Resin from Sigma.

^a Protein description in Swiss-Prot database.

^b Accession number was recorded as a reference for the identification in Swiss-Prot database.

^c It was based on the annotation in Swiss-Prot database.

In fact, the mass window of 0.002 Da could also be achieved with microTOF. However, based on our result, the accuracy of quantification was affected for less than 2% if 0.02 Da was used, but it reduced the labor intensity of selecting hrEICs significantly. As shown in Fig. 4, the hrEICs of one identified peptide of HRP in both samples were generated, and their peak areas were used for relative quantification.

To evaluate the accuracy of the label-free quantification, 2 μ l and 3 μ l of the low-abundance glycoproteins obtained in our strategy were loaded and analyzed with microTOF, respectively. Both the MS and MS/MS analyses were conducted in triplicate. Because the ionization efficiencies of 2 μ l and 3 μ l runs were highly comparable, we used them in testing the accuracy of relative quantification by microTOF without normalization with HRP. In Table 3, 20 proteins were quantified using the peak area values of multiple identified peptides of each protein, and the data for the rest proteins were shown in Supplemental Table 1. As shown in Table 3, the quantification error of 13 proteins (65.00%) was less than 4%; that of 5 proteins (25.00%) was between 4% and 10%; that of 2 proteins (10.00%) was between 10% and 20%. It indicated that the quantification of microQTOF was very accurate compared with other quantification methods [27]. The commonly known relatively higher-abundance proteins, such as ceruloplasmin, hemopexin and the added HRP, except for top 12 proteins were quantified in Table 3. It was obvious that the quantification accuracy of them was consistent with others. Thus, the concentration of all of the proteins was within the linear range of quantification in our experiment. The relative standard deviation (RSD) of the quantification values calculated from different peptides of the same protein was less than 30%, and 17 of 20 (85%) was even less than 20%. The RSD of the triplicate analyses for the quantification of proteins was less than 15%, and 11 of 20 (55%) was even less than 5.00%, which indicated that the MS runs were highly reproducible. We then evaluated the accuracy of quantification if only one peptide of each protein was used for quantification. Statistically, for 60 peptides used for quantification in Table 3, the quantification error of 52 of them (86.67%) was less than 30%; that of 6 of them (10.00%) was between 30% and 50%; that of 2 of them (3.33%) was between 50% and 100%. It indicated that even if only one peptide was identified for one protein, it would be safely used for relative quantification.

3.3. Evaluation of the possible influence of immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization on protein quantification

Because of the high complexity and large dynamic range of human serum proteins, it was essential to separate low-abundance proteins before quantification could be conducted [28,29]. Isolation of glycoproteins with lectins was also widely used for separation of them from other high-abundance proteins in the analysis of glycoproteins. The efficiency, reproducibility and recovery of the immunodepletion and lectin enrichment have been evaluated by many researchers [2,9,15]. However, no research has been done on whether the combined use of them would result in the significant change of protein quantification. Because our aim was to quantify low-abundance glycoproteins between normal and diseased samples, we must ascertain that the above-mentioned procedures would not result in the significant change of protein relative concentration.

In order to evaluate and eliminate the possible influence of immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization on protein quantification, 0.2% and 0.4% HRP were added into two aliquots of the same serum sample at the beginning. Subsequently, they were subjected to the same treatments until MS/MS analysis was finished. Both the sample preparation and MS analysis were conducted in duplicate. And then

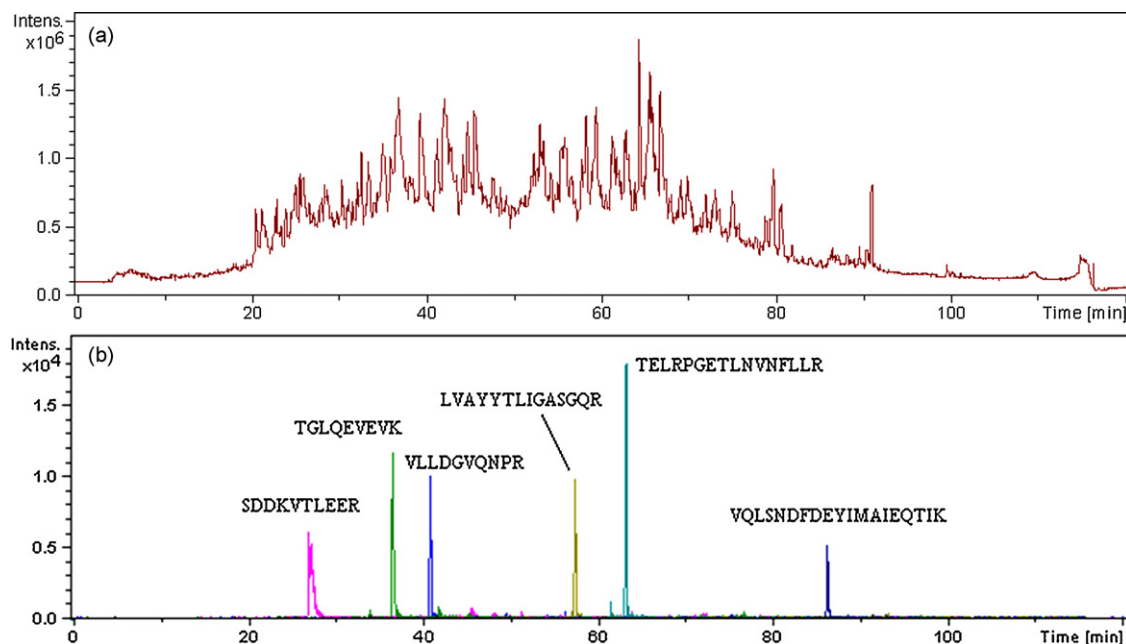


Fig. 3. Total ion chromatogram (TIC) (a), and hrEICs of identified peptides (b) from human Complement C3 precursor. Because of the high mass accuracy, it was quite easy to distinguish the hrEICs from the background.

the relative ratio of 20 proteins between the two samples was quantified in Table 4, and the data for the rest proteins were shown in Supplemental Table 2. As shown in Table 4, the quantification error of 6 proteins (30%) was less than 4%; that of 5 proteins (25%) was between 4% and 10%; that of 8 proteins (40%) was between 10% and 30%; only one of them (5%) was 32.86%. It indicated that the procedures presented in our strategy had an acceptable impact on the relative quantification of the added HRP and other enriched low-abundance glycoproteins. To demonstrate the repeatability of the whole procedure, we determined the normalized ratio of peak area for 60 individual peptides from different proteins between different sample preparations for the same sample. An average value of

96.53% was obtained, and the RSD was 23.62%. The linear correlation of the peak intensities between the two preparations was also plotted in Supplemental Figure 1. It indicated that the sample preparations were highly reproducible. We also determined the concentration of proteins obtained in each step of our strategy by Bradford assay, and it also indicated that each step was consistent among different runs (data not shown). Since there were many steps that might affect the accuracy of quantification in our experiment, we used the quantification data for monitoring the repeatability of the whole procedure. As shown in Table 4, the RSD of averages obtained from duplicate sample preparation and duplicate MS runs (four repetitions in total) was very low, which indicated that the

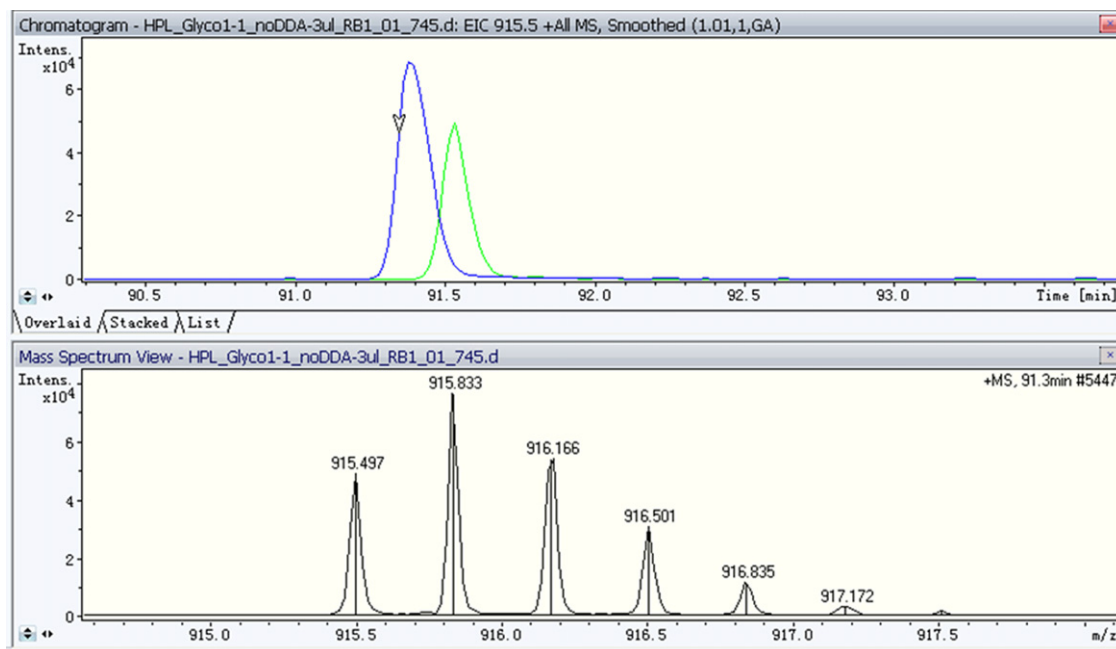


Fig. 4. The hrEICs of one identified peptide of HRP in both samples used for relative quantification. There were no other background peaks that affected the accuracy of relative quantification.

Table 3The relative quantification of 20 identified glycoproteins between two completely same samples in which 2 μ l and 3 μ l samples were loaded, respectively.

No.	Protein description ^a	Number of peptides for quantification	Average ^b	RSD% ^c of different peptides	RSD% ^d of triplicate analyses	Deviation from theoretical value ^e
1	HRP	5	1.51	10.85%	1.38%	0.89%
2	Ceruloplasmin	3	1.63	13.57%	4.91%	8.89%
3	Hemopexin precursor	3	1.48	12.28%	2.44%	1.33%
4	ITI heavy chain H4	3	1.55	24.66%	3.10%	3.42%
5	Antithrombin III	3	1.55	8.12%	2.82%	3.64%
6	Complement factor B	3	1.49	12.84%	6.93%	0.40%
7	Heparin cofactor 2	3	1.56	12.96%	4.50%	3.89%
8	Complement factor I	3	1.63	2.65%	7.30%	8.51%
9	Lumican	3	1.70	13.04%	4.71%	13.33%
10	Angiotensinogen	3	1.60	10.33%	9.57%	6.78%
11	Alpha-2-HS-glycoprotein	3	1.51	15.52%	10.32%	0.36%
12	Complement C4-A	3	1.54	17.81%	2.01%	2.91%
13	Kininogen-1 precursor	5	1.56	23.39%	2.37%	4.09%
14	AMBP protein	3	1.45	13.70%	5.67%	3.38%
15	Complement C9	3	1.51	18.61%	1.96%	0.84%
16	ITI heavy chain H1	3	1.73	11.48%	7.85%	15.24%
17	Alpha-2-antiplasmin	3	1.61	15.31%	5.03%	7.60%
18	Complement C2	2	1.54	17.81%	9.30%	2.51%
19	PGRP-L	2	1.45	7.56%	11.97%	3.27%
20	Prothrombin	1	1.55	NA	4.36%	3.22%
	Total	60				

^a Protein description in Swiss-Prot database. All analyses were performed in triplicate.^b The relative quantification value from each individual peptide was calculated respectively, and then the average of them was regarded as the ratio of the protein. The average of triplicate analyses was regarded as the final ratio.^c Relative standard deviation (RSD%) of the quantification value was calculated from different peptides of the same protein.^d Relative standard deviation (RSD%) of the triplicate analyses.^e The theoretical value is equal to 1.5.

overall procedure was highly reproducible. Furthermore, the internal standard added at the beginning was also helpful in reducing the impact of some occasional factors happened during the whole procedure.

In our work, we aimed to establish the reasonable strategy for enriching and quantifying low-abundance glycoproteins between different samples. Only two lectins were used in our experi-

ment because it was easier to conduct the experiment with less lectins. If necessary, you could increase the number of lectins in one M-LAC column to 8 [25]. We could also use ProteoPrep 20 Immunoaffinity Depletion Kit from Sigma in the depletion of high-abundance proteins if the lower-abundance proteins were required for analysis. It would be also applicable to the conclusion obtained in this work. Furthermore, only 62 low-abundance

Table 4

The relative quantification of 20 identified glycoproteins between two same serum samples with HRP added at the ratio of 1:2.

No.	Protein description ^a	Number of peptides for quantification	Average ^b	RSD% ^c of different peptides	RSD% ^d of quadruplicate analyses	Deviation from theoretical value ^e
1	HRP	5	0.47	13.5%	18.32%	6.00%
2	Ceruloplasmin	7	2.02	22.18%	11.81%	1.00%
3	hemopexin precursor	8	2.43	13.90%	6.51%	21.5%
4	Complement factor B precursor	5	2.18	5.45%	4.97%	9.00%
5	Antithrombin III	4	2.05	19.10%	4.56%	2.35%
6	Kininogen-1 precursor	5	2.32	19.15%	8.59%	15.87%
7	Plasma protease C1 inhibitor precursor	3	2.45	39.90%	10.21%	22.47%
8	Angiotensinogen	2	1.81	16.64%	17.83%	9.26%
9	Beta-2-glycoprotein 1	2	1.92	32.90%	8.67%	4.03%
10	Complement factor H precursor	2	2.66	48.70%	23.19%	32.86%
11	Alpha-2-HS-glycoprotein	3	1.64	17.38%	26.89%	18.00%
12	Complement C4-A	3	2.06	4.34%	20.55%	3.00%
13	ITI heavy chain H4	3	2.03	9.96%	11.39%	1.50%
14	AMBP protein	3	2.43	22.44%	7.65%	21.50%
15	Complement C9	3	1.99	24.26%	4.77%	0.50%
16	ITI heavy chain H1	3	2.27	27.71%	28.52%	13.50%
17	Alpha-2-antiplasmin	3	1.96	15.56%	24.87%	2.00%
18	Complement C2	2	2.51	5.63%	3.87%	25.50%
19	PGRP-L	2	2.15	3.35%	5.01%	7.50%
20	Prothrombin	1	2.21	NA	10.45%	10.50%
	Total	69				

^a Protein description in Swiss-Prot database. All analyses were performed in quadruplicate.^b The relative quantification value from each individual peptide was calculated respectively, and then the average of them was regarded as the ratio of the protein. The average of quadruplicate analyses (duplicate sample preparation and duplicate MS runs) was regarded as the final ratio.^c Relative standard deviation (RSD%) of the quantification value calculated from different peptides of the same protein.^d Relative standard deviation (RSD%) of the quadruplicate analyses.^e The theoretical value of the ratio of HRP is equal to 0.5, while that of other proteins was 2.0 after normalization with HRP.

proteins were identified in our experiment although the quantification of them would be very accurate. In the future, if MS equipment with comparable resolving power and mass accuracy but higher scanning speed was utilized, it was sure that the number of identified proteins would increase greatly. However, by comparing our results with that of others, most of them were identified as low-abundance glycoproteins. The most low-abundance glycoproteins quantified in our experiment were at the concentration of 1–20 $\mu\text{g}/\text{ml}$ in normal human serum, such as Complement component C6, PEDF, and so on [30,31]. It could be regarded as the detection limit of our method for serum samples. Generally, it was impossible to detect these proteins directly from human serum by LC-MS/MS. Some of them might be regarded as biomarker candidates or be used for monitoring the postoperative status of patients if concentration changes of them were found between different samples.

4. Conclusions

Here, we described the combined strategy of immunodepletion, enrichment and quantification of low-abundance glycoproteins in human serum. Based on our results, Seppro IgY12 columns were specific, efficient and reproducible in removing top 12 proteins; micrOTOF-Q from Bruker Daltonics could be efficiently used in the identification and quantitative analysis of complex samples, such as human serum. Furthermore, the procedures presented in this work, such as immunodepletion, enrichment of glycoproteins, trypsin digestion and peptide ionization, would not affect the relative quantification of the enriched low-abundance glycoproteins significantly. The label-free quantification method in this paper can be effectively applied to the screening of low-abundance glycoprotein biomarker candidates and the monitoring of the postoperative status of patients. In the future, the use of more sophisticated MS equipment would extend the application of our strategy.

Acknowledgements

The authors thank Bruker Daltonics for instrumental support. We also thank Zhensheng Xie and Prof. Fuquan Yang from Institute of Biophysics, Chinese Academy of Sciences, for providing the convenience of packing C18 columns.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2004.08.011.

References

- [1] W. Morelle, K. Canis, F. Chirat, V. Faid, J.C. Michalski, *Proteomics* 6 (2006) 3993.
- [2] K. Ueda, T. Katagiri, T. Shimada, S. IrieS, T.A. Sato, Y. Nakamura, Y. Daigo, *J. Proteome Res.* 6 (2007) 3475.
- [3] Y. Ren, J. Wang, J. Xia, C. Jiang, K. Zhao, R. Li, N. Xu, Y. Xu, S. Liu, *J. Proteome Res.* 6 (2007) 2812.
- [4] T. Fujimura, Y. Shinohara, B. Tissot, P.C. Pang, et al., *Int. J. Cancer* 122 (2008) 39.
- [5] A. Vanderver, R. Schiffmann, M. Timmons, K.A. Kellersberger, D. Fabris, E.P. Hoffman, J. Maletkovic, Y. Hathout, *Clin. Chem.* 51 (2005) 2031.
- [6] W.C. Cho, *Mol. Cancer* 6 (2007) 25.
- [7] S.K. Chatterjee, B.R. Zetter, *Future Oncol.* 1 (2005) 37.
- [8] J. Zhao, D.M. Simeone, D. Heidt, M.A. Anderson, D.M. Lubman, *J. Proteome Res.* 5 (2006) 1792.
- [9] J. Martosella, N. Zolotarjova, H. Liu, G. Nicol, B.E. Boyes, *J. Proteome Res.* 4 (2005) 1522.
- [10] H.L. Huang, T. Stasyk, S. Morandell, et al., *Electrophoresis* 27 (2006) 1641.
- [11] E.I. Goufman, S.A. Moshkovskii, O.V. Tikhonova, et al., *Biochemistry (Mosc.)* 71 (2006) 354.
- [12] P. Dowling, L. O'Driscoll, P. Meleady, et al., *Electrophoresis* 28 (2007) 4302.
- [13] K.H. Yu, A.K. Rustgi, I.A. Blair, *J. Proteome Res.* 4 (2005) 1742.
- [14] Y. Wang, S.L. Wu, W.S. Hancock, *Glycobiology* 16 (2006) 514.
- [15] Z. Yang, W.S. Hancock, *J. Chromatogr. A* 1053 (2004) 79.
- [16] S.L. Wu, H. Amato, R. Biringer, G. Choudhary, P. Shieh, W.S. Hancock, *J. Proteome Res.* 1 (2002) 459.
- [17] Z. Yang, L.E. Harris, D.E. Palmer-Toy, W.S. Hancock, *Clin. Chem.* 52 (2006) 1897.
- [18] L.A. Echan, H.Y. Tang, N. Ali-Khan, K. Lee, D.W. Speicher, *Proteomics* 5 (2005) 3292.
- [19] Y. Gong, X. Li, B. Yang, W. Ying, D. Li, Y. Zhang, S. Dai, Y. Cai, J. Wang, F. He, X. Qian, *J. Proteome Res.* 5 (2006) 1379.
- [20] T. Linke, S. Doraiswamy, E.H. Harrison, *J. Chromatogr. B* 849 (2007) 273.
- [21] A.V. Rapkiewicz, V. Espina, E.F. Petricoin 3rd, L.A. Liotta, *Eur. J. Cancer* 40 (2004) 2604.
- [22] I.C. Guerrero, O. Kleiner, *Biosci. Rep.* 25 (2005) 71.
- [23] A. Nita-Lazar, H. Saito-Benz, F.M. White, *Proteomics* 8 (2008) 4433.
- [24] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, *Anal. Bioanal. Chem.* 389 (2007) 1017.
- [25] S.H. Heo, S.J. Lee, H.M. Ryoo, J.Y. Park, J.Y. Cho, *Proteomics* 7 (2007) 4292.
- [26] M.J. Ford, J.L. Cantone, C. Polson, J.H. Toyn, J.E. Meredith, D.M. Drexler, *J. Neurosci. Methods* 168 (2008) 465.
- [27] T. Chris, PRG2006 Research Study – Relative Quantification, <http://www.abrf.org/prg>.
- [28] M. Ramström, C. Hagman, J.K. Mitchell, P.J. Derrick, P. Håkansson, J. Bergquist, *J. Proteome Res.* 4 (2005) 410.
- [29] N.I. Govorukhina, T.H. Reijmers, S.O. Nyangoma, A.G. van der Zee, R.C. Jansen, R. Bischoff, *J. Chromatogr. A* 1120 (2006) 142.
- [30] F. Tedesco, P.J. Lachmann, *Clin. Exp. Immunol.* 9 (1971) 359.
- [31] S.V. Petersen, Z. Valnickova, J.J. Enghild, *Biochem. J.* 374 (2003) 199.