Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 845 (2007) 121–137
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

High-sensitivity profiling of glycoproteins from human blood serum through multiple-lectin affinity chromatography and liquid chromatography/tandem mass spectrometry

Milan Madera, Yehia Mechref, Iveta Klouckova, Milos V. Novotny *

Department of Chemistry, Indiana University, 800 E Kirkwood Ave., Bloomington, IN 47405, United States

Received 13 April 2006; accepted 28 July 2006
Available online 20 September 2006

Abstract

We report here the use of high-performance lectin affinity enrichment of glycoproteins at microscale levels using a series of silica-bound lectins. The potential of this approach is being demonstrated for the glycoprotein enrichment from microliter volumes of human blood serum. Individual injections of sample to the affinity microcolumns packed with four lectin materials with different glycan specificities (Con A, SNA-I, UEA-I, PHA-L), followed by off-line reversed-phase pre-fractionation and nano-LC/MS/MS, permitted identification of 108 proteins in the lectin-bound fractions spanning a concentration dynamic range of 7–10 orders of magnitude. In contrast, multi-lectin microcolumn affinity chromatography, an alternative enrichment approach allowed identification of only 67 proteins. An attractive feature of high-performance lectin affinity chromatography at microscale levels is the substantial reduction of sample losses that are commonly experienced with extensive sample preparation needed for larger sample volumes.

Keywords: Lectin; Multiple lectin; Affinity chromatography; Glycoproteins; Mass spectrometry; Human blood serum; High sensitivity

1. Introduction

Posttranslational modifications (PTM) of proteins are among the key biological regulators of function, activity, localization, and interaction. The fact that no more than 30,000–50,000 proteins are encoded by the human genome underlines the importance of posttranslational modifications in modulating the activities and functions of proteins in health and disease [1]. More than 100 different types of PTM have thus far been documented, each playing a different role [2]. These PTMs can be associated with a single or multiple sites on polypeptides. The overall proteome complexity is, accordingly, substantially augmented by the diversity of PTM, adding methodological complexity to the field of proteomics. A major challenge to the field is to define the characteristics and dynamics of PTMs in cells, tissues and organisms.

With approximately 50% of all proteins now considered to be glycosylated [3], this type of PTM has been widespread and physiologically important in mammalian systems. A growing list of glycoproteins has been shown to act through a recognition of oligosaccharide chains and their microheterogeneities at the site of modification. For example, the importance of glycans as biological determinants has been reflected in the fact that cell–cell interactions involve sugar–sugar or sugar–protein specific recognition. Consequently, aberrant glycosylation has now been recognized as an attribute of many mammalian diseases, including hereditary disorders, immune deficiencies, neurodegenerative diseases, cardiovascular conditions, and cancer [4,5]. As many potential disease biomarkers may be glycoproteins present in only minute quantities in tissue extracts and physiological fluids, glycoprotein isolation and enrichment steps prior to analysis can be extremely helpful in a search for such biomarkers.

Numerous attempts have been made [6] to develop enrichment methods for glycoproteins from complex biological samples. Logically, the great majority of these enrichment methodologies rely on the use of immobilized lectins, which in their more modern versions permit a more or less selective enrichment of the pools of glycoproteins for proteomic/glycomic studies [7–10]. From a large number of lectins, of either plant

* Corresponding author. Tel.: +1 812 855 4532; fax: +1 812 855 8300.
E-mail address: novotny@indiana.edu (M.V. Novotny).

1570-0232/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2006.07.067

or animal origin, available for a biospecific interaction [11], relatively few have been immobilized to solid supports. During a common use of lectin affinity chromatography as a glycoprotein isolation tool, a complex mixture of proteins is applied to a lectin immobilized on a suitable solid support, while the unbound proteins can be washed out with a binding buffer and the interacting glycoproteins are subsequently displaced by washing with an elution buffer, which usually contains a hapten sugar. Lectin affinity chromatography can be conducted in different formats: tubes [8]; packed columns [9]; microfluidic channels [12]; or specialty solid surfaces, such as colloidal gold [13] and affinity membranes [14].

While lectin affinity chromatography in its glycoprotein isolation mode has now become a common methodology in numerous biochemical laboratories, its use in the analytical/quantitative mode has been less common. The analytical mode is likely to find an increasing demand in contemporary proteomic investigations where differential measurements are needed on a multitude of glycoproteins, within an extensive concentration range. Simultaneously, development of high-throughput analytical systems is likely to necessitate lectin enrichment devices that operate under the conditions of high pressure. Additionally, the emphasis on analyzing small samples without the losses of trace analytes in the system distinctly favors miniaturization of all system's components, including the lectin preconcentration step. Although there have been promising attempts to miniaturize lectin affinity devices through surface-immobilized lectin arrays [15] or lectin microcolumns coupled to reversed-phase liquid chromatography [16] and liquid chromatography/mass spectrometry [17,18], a comprehensive enrichment of glycoproteins from complex protein mixtures has not been systematically addressed. While the recent studies using single immobilized lectins or multi-lectin affinity chromatography [19,20] resulted in identification of more than 150 glycoproteins present in human blood serum [10], there is still a considerable room for improvement to fulfill the demands of comprehensive proteomics.

Here we report the use of sequential lectin affinity chromatography in silica-based microcolumns [17] at a substantially reduced sample volume than commonly practiced. In profiling human serum glycoproteins, we have investigated four different silica-immobilized lectins for their specificities and the capability to selectively enrich different glycoproteins within a substantial dynamic concentration range. The results suggest that a sequential use of lectin microcolumns is a superior enrichment procedure to the previously used multi-lectin affinity approach. As shown below, from only a 20-μL sample aliquot, we were able to identify 108 glycosylated proteins through coupling this enrichment technique with capillary liquid chromatography (LC) and tandem mass spectrometry (MS/MS).

2. Materials and methods

2.1. Chemicals

Silica NUCLEOSIL 10-1000 (10 μm particle size, 1000 Å pore size) was purchased from Macherey-Nagel Inc. (Easton, PA), while all lectins were obtained from EY Laboratories (San Mateo, CA). Dithiothreitol (DTT) and iodoacetamide (IAA) were purchased form Bio-Rad (Hercules, CA), while the sequencing-grade TPCK-treated trypsin originated from Sigma–Aldrich (St. Louis, MO). Human serum samples pooled from healthy male donors and all other chemicals were also received from Sigma–Aldrich. Buffers and other stock solutions were prepared by dissolving of the appropriate materials in Millipore deionized water (Billerica, MA). All samples were stored at −20 °C and used immediately after thawing.

2.2. Preparation of lectin-silica

Derivatization of the macroporous silica and the immobilization of the lectins was performed according to our recently published procedure [17] which is a modification of a previously published method [18,21]. Briefly, a 1-g aliquot of NUCLEOSIL 10-1000 resin was sequentially washed with water, 6 M hydrochloric acid, and water again. The resins were then dried at 150 °C overnight. After resuspension in 15 mL toluene dried with a molecular sieve, 200 μL of 3-glycidoxypropyltrimethoxysilane was added, followed by 5 μL of triethylamine. Next, the suspension was placed under reflux and stirred at 105 °C for 16 h. The resulting epoxy-silica was washed with toluene, acetone, and ether prior to drying under vacuum. Next, a 1-g aliquot of epoxy-silica was resuspended in 100 mL of 20 mM sulfuric acid and hydrolyzed to diol-silica at 90 °C for 3 h. The resulting diol-silica resins were then washed with water, ethanol, and ether and dried under reduced pressure. A 1-g aliquot of diol-silica was mixed with 20 mL of acetic acid/water (90/10) prior to the addition of a 1-g aliquot of sodium periodate. The mixture was then stirred at room temperature for 2 h to oxidize diols to aldehydes. Aldehyde-silica resins were then washed with water, ethanol, and ether and dried under reduced pressure.

Silica-bound lectin resins were prepared individually. A 10-mg aliquot of *Canavalia ensiformis* lectin (Con A), 5 mg of *Sambucus nigra* lectin (SNA-I), 5 mg of *Ulex europaeus* lectin (UEA-I) or 5 mg of *Phaseolus vulgaris* lectin (PHA-L) were first suspended in 1 mL of 0.1 M sodium bicarbonate containing 0.5 M sodium chloride and 10 mg of sodium cyanoborohydride. Each solution was then added to 125 mg of the prepared aldehyde-silica, and the suspension was vortexed, deaerated for 5 min, and stirred at room temperature for 3 h. A 5-mg aliquot of sodium borohydride was added in portions and stirred for another hour to reduce excessive aldehyde groups to diols. All lectin-silicas were then washed with water, 0.1 M sodium bicarbonate with 0.5 M sodium chloride, and water again. Con A-silica was washed and resuspended in 10 mM Tris–HCl (pH 7.4) containing 150 mM NaCl, 1 mM $MnCl_2$, 1 mM $MgCl_2$, 1 mM $CaCl_2$ and 0.02% NaN_3 while other lectin-silicas were washed and resuspended in 50 mM sodium PBS (pH 7.4) containing 150 mM NaCl and 0.02% NaN_3. Immobilized lectins were stored at 4 °C and pre-warmed to room temperature prior to use.

The amount of lectin attached to silica was determined by the Bradford assay [22] from the difference of the lectin amount present in coupling solution before and after immobilization.

Table 1
Binding and elution buffers, and immobilized lectins used for the sequential and multi-lectin affinity modes employed in this work

Lectin (source)	Binding buffer	Elution buffer
Con A (*Canavalia ensiformis*)	10 mM Tris–HCl (pH 7.4), 150 mM NaCl, 1 mM $MnCl_2$, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 0.02% NaN_3	0.1 M methyl-α-D-mannoside in binding buffer
SNA-I (*Sambucus nigra*)	10 mM sodium phosphate buffer (pH 7.4), 150 mM NaCl, 0.02% NaN_3	0.1 M D-lactose in binding buffer
UEA-I (*Ulex europaeus*)	10 mM sodium phosphate buffer (pH 7.4), 150 mM NaCl, 0.02% NaN_3	0.1 M α-L-fucose in binding buffer
PHA-L (*Phaseolus vulgaris*)	10 mM sodium phosphate buffer (pH 7.4), 150 mM NaCl, 0.02% NaN_3	0.4 M *N*-acetyl-D-glucosamine in binding buffer
Multi-lectin	10 mM sodium phosphate buffer (pH 7.4), 150 mM NaCl, 0.02% NaN_3	Binding buffer containing 0.1 M methyl-α-D-mannoside, 0.1 M D-lactose, 0.1 M α-L-fucose and 0.4 M *N*-acetyl-D-glucosamine

This amount was determined to be 50 mg of Con A/g silica and 35 mg of all other lectins/g silica.

2.3. Packing of columns

Lectin-silica slurry was packed into 500 μm × 5 cm PEEK tubing with the unions containing 0.5-μm stainless steel frits (Scivex, Upchurch Scientific, WA), using Capillary Perfusion Toolkit (Perseptive Biosystems, Framingham, MA) and a high-pressure liquid pump (Beckman, Fullerton, CA) operating at 2500 psi. After packing, lectin microcolumns were washed with the appropriate binding buffer (see Table 1), at 5 μL/min flow rate for 20 min, and stored at 4 °C. These columns were pre-conditioned with binding buffer at ambient temperature prior to use.

A multi-lectin affinity microcolumn was prepared by packing a slurry, consisting of the evenly mixed amounts of the individual lectin-silicas into a 1 mm × 15 cm PEEK tubing with the unions and stainless-steel frits, using the same packing procedure. This microcolumn was washed with 50 mM sodium PBS (pH 7.4) containing 150 mM NaCl and 0.02% NaN_3 at 40 μL/min flow rate for 20 min, and stored at 4 °C. The microcolumn was preconditioned with the binding buffer (Table 1) at ambient temperature prior to use.

A C18 microcolumn was prepared by packing C18 silica slurry (5 μm, 300 Å) into 0.5 mm × 5 cm PEEK tubing using the same procedure as that for the packing of sequential-lectin affinity chromatography microcolumns with methanol as a packing solvent.

The C18 nanocolumn utilized for nano-LC-ESI/MS/MS runs was prepared by packing silica slurry (Magic C18, Waters, Manford, MA) into a 75-μm i.d. pulled-tip fused silica capillary (Polymicro Technologies, Phoenix, AZ) using a pressurized gas and, finally, cut to its 15-cm length. The column was washed with 80% acetonitrile mobile phase containing 0.1% formic acid and, subsequently, equilibrated with 3% acetonitrile with 0.1% formic acid at 250 nL/min flow rate.

2.4. Affinity enrichement on individual lectin microcolumns

All experiments were conducted using a syringe pump (KD Scientific, Hoilliston, MA) connected to a Rheodyne injection valve (Rohent Park, CA) with a 25-μL sample loop. The different lectin microcolumns employed in this study dictated the use of different binding and elution buffers, as summarized in Table 1. A 5-μL aliquot of human blood serum was diluted five-fold with an appropriate lectin binding buffer prior to injection. Unbound proteins were washed from the lectin microcolumn with a binding buffer at 5 μL/min flow rate for 30 min and collected into an Eppendorf tube. Next, bound proteins/glycoproteins were eluted using the appropriate elution buffer (see Table 1) at 5 μL/min flow rate for 30 min.

2.5. Multi-lectin affinity chromatography

Multi-lectin affinity microcolumn was connected to an Äkta Purifier instrument (Amersham Biosciences, NJ). A 20-μL aliquot of human blood serum was mixed with an equal volume of the lectin binding buffer (see Table 1), filtered through 0.5 μm filter prior to injection. Unbound proteins were washed from the lectin column with the binding buffer at 40 μL/min flow, while a 1.5 mL aliquot of the unbound fraction was collected. Bound proteins were displaced from the column with a mixed elution buffer, and a 1.5 mL aliquot of bound fraction was collected.

2.6. Sample clean-up and tryptic digestion prior to LC–MS/MS analysis

Collected unbound and bound fractions were desalted on MICROCON 10 kDa cut-off membrane filters (Millipore, Billerica, MA), lyophilized and resuspended in a reducing solution, consisting of 6 M guanidine hydrochloride and 20 mM DTT prepared in 50 mM ammonium bicarbonate buffer solution. Samples were then incubated at 65 °C for 1 h prior to alkylation through the addition of 20 μL of 0.75 M iodoacetamide, followed by incubation at room temperature in the dark for 40 min. Reduced and alkylated samples were centrifuged, exchanged into 50 mM ammonium bicarbonate (pH ∼ 8.0) using MICROCON 10 kDa cut-off membrane filters, and incubated with trypsin (1%, w/w) at 37 °C for 18 h.

2.7. Reversed-phase fractionation of tryptic digests

The pH of a 200-μL aliquot of the tryptic digest was adjusted through the addition of 0.4 μL of TFA prior to loading on a C18 microcolumn, pre-conditioned with 80% acetonitrile and equilibrated with a washing solvent (3% acetonitrile containing 0.1% TFA). The column was then washed with the same solvent at 5 μL/min flow for 20 min. Next, peptides were eluted, stepwise,

using 10, 20, 30, 50 and 80% acetonitrile in 0.1% TFA, separately collected, dried in a SpeedVac unit, and resuspended in the appropriate amount of deionized water prior to LC–MS/MS analysis.

2.8. LC–MS/MS analysis of digested proteins and data evaluation

Tryptic digests were analyzed by capillary reversed-phase liquid chromatography connected to an LCQ-DECA XP ion-trap mass spectrometer (Thermo Finnigan, San Jose, CA). Peptides were separated through a linear gradient, from 3 to 55% acetonitrile, 0.1% formic acid over 55 min. Mobile-phase composition was then changed to 80% acetonitrile over 10 min, thus eluting highly hydrophobic peptides. Column eluent was directly electrosprayed into the mass spectrometer using a 3-kV spraying voltage. Ions were scanned and detected within the *m*/*z* range from 300 to 1600.

Analysis of the data files generated by a mass spectrometer usually requires further processing, filtering and converting into the format suitable for a database search. Despite the common use of a sophisticated software, such as SEQUEST or Mascot, identification of the proteins from peptide sequences generated from MS/MS data is always accompanied by either false-positive or false-negative assignments. Even though the search engines provide quite sufficient probability of positive matches (>95%), default-filtering thresholds must often be adjusted to eliminate false-negatives while keeping the credibility of the positive hits. Created DTA files were subjected to Mascot search against SwissProt database, allowing the processing of all tryptic and semi-tryptic peptides up to 2 miscleavages. The searching criteria included 2+ and 3+ charged ions, while 1+ ions were not considered relevant because tryptic peptides should have at least two easily protonizable sites, N-terminus and side-chain of lysine or arginine on C-terminus [23]. The data were then "filtered" with an in-house parser, with Mowse probability score threshold set to 30, thus ensuring that the probability of the peptide match to be a random event is less than 5%. Although this is considered as fairly confident, there is still a possibility of having some false positives. Therefore, the level of confidence was increased by applying additional filtering criteria. The peptides with sequences containing less than six amino acids and/or peptides smaller than 600 Da were considered false positives, since the absolute majority of such small peptides do not appear to be correctly assigned. Also, tryptic peptides containing internal KK, KR, KR or RR motifs were excluded because trypsin would have efficiently cleaved at least one of the two bonds [24].

3. Results and discussion

High-performance (rigid) LC packings used in LC offer some advantages over traditional materials: sharp elution profiles, better efficiencies and decreased analysis time [17,25]. The use of pressure-compatible small silica particles packed into concentrating affinity microcolumns thus appears attractive in developing comprehensive, automatable analytical systems for protein enrichment and affinity-based fractionations in proteomic research and clinical applications. Miniaturization is also important when working with small amounts and volumes of biological material, as it reduces significantly sample losses due to a non-specific adsorption on the affinity media or the walls of separation channels.

To demonstrate the advantages of small-scale sample enrichment, we have used here sequential-lectin affinity microcolumns to investigate glycoproteins derived from human blood serum. All the lectins used in this work were first immobilized on derivatized macroporous silica which offers excellent mechanical stability, permeability and adequate loading capacity (ranging from 35 to 50 mg lectin/g silica). Due to the expectation of a wide range of glycan structures that could be attached to different human serum glycoproteins, four lectin materials with different specificities were prepared to ensure a maximum enrichment. First, ConA is a lectin with a broad specificity, interacting mainly with mannose, glucose and galactose residues, and capable of recognizing certain *N*-glycans with the chitobiose core [26]. Conversely, SNA-I is a lectin with a very narrow specificity, targeting primarily α(2,6)-linked sialylated structures [27]. Since a majority of all glycoproteins present in eukaryotes are likely to be sialylated, using an appropriate set of lectins together should allow a substantial coverage of human glycoproteome after enrichment. Alternatively, a targeted enrichment might work beneficially to single out the biomarkers with certain known glycosylation attributes that are characteristic of disease conditions. For example, the extent of fucosylation can be monitored through the use of UEA-1 lectin which interacts more strongly with α(1,2)-linked fucose, as compared to weaker interactions with α(1,3) and α(1,6) linked fucose structural entities [28]. The PHA-L lectin distinctly recognizes the Galβ(1,4)GalNAcβ(1,2)Man trisaccharide attachment, but, in some cases, the presence of a terminal α(2,6)-linked sialic acid will prevent a lectin binding [29]. The PHA-L lectin is likely to be a valuable enrichment tool due to its binding preferences for the glycan structures which have been implicated in various stages of cancer [30].

Since the primary objective of this study has been the development of an approach allowing a comprehensive enrichment of glycoproteins from any biological sample, the sequential arrangement of different lectin microcolumns was deemed desirable. As seen in Table 2, this analytical strategy permitted us to identify 108 proteins from a small-volume sample. It should be noted that a comparable number of proteins is typically seen with the use of lectin affinity chromatography with sample volumes around 1 mL [19,20,31], for example, through the use of multi-lectin enrichment [19] through mixing the beads with different lectins.

The enhanced affinity of proteins toward the individual lectins is expressed as the number of derived peptides, which were identified in a particular fraction. It is noticeable that 92 proteins (over 85% of the total identified number) feature the *N*-glycosylation (NXS/T) motif. Still, the remaining proteins may or may not be glycosylated, as their presence in the lectin-bound fractions could potentially be explained by either (1) some *O*-glycosylation which does not feature an easily discernable motif; (2) a non-specific binding; (3) association of some non-

Table 2
List of proteins identified in the lectin-bound fractions after high-performance affinity enrichment of human blood serum (individual lectins) at microscale levels

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	Number of peptides matched			
						Con A	SNA-I	UEA-I	PHA-L
1	A1AG1_HUMAN	(P02763) Alpha-1-acid glycoprotein 1 precursor (AGP 1) (orosomucoid 1) (OMD 1)	23725	4.93	5	0	2	1	5
2	A1AG2_HUMAN	(P19652) Alpha-1-acid glycoprotein 2 precursor (AGP 2) (orosomucoid 2) (OMD 2)	23873	5.03	5	1[a]	1[a]	0	3
3	A1AT_HUMAN	(P01009) Alpha-1-antitrypsin precursor (alpha-1 protease inhibitor) (alpha-1-antiproteinase)	46878	5.37	3	4	7	4	3
4	A2MG_HUMAN	(P01023) Alpha-2-macroglobulin precursor (alpha-2-M)	164600	6.00	8	27	9	28	18
5	AACT_HUMAN	(P01011) Alpha-1-antichymotrypsin precursor (ACT)	47792	5.33	6	4	4	2	4
6	ADT2_HUMAN	(P05141) ADP; ATP carrier protein; fibroblast isoform (ADP/ATP translocase 2)	32971	9.76	1	0	0	0	1
7	AKIB1_HUMAN	(Q9P2G1) Ankyrin repeat and IBR domain containing protein 1 (fragment)	137695	5.19	4	0	0	1[a]	1[a]
8	ALBU_HUMAN	(P02768) Serum albumin precursor	71317	5.92	0	24	28	32	37
9	ANT3_HUMAN	(P01008) Antithrombin-III precursor (ATIII) (PRO0309)	53025	6.32	4	0	0	0	2
10	APOA1_HUMAN	(P02647) Apolipoprotein A-I precursor (Apo-AI) (ApoA-I)	30759	5.56	0	9	5	6	9
11	APOA2_HUMAN	(P02652) Apolipoprotein A-II precursor (Apo-AII) (ApoA-II)	11282	6.27	0	4	2	0	1
12	APOA4_HUMAN	(P06727) Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)	45343	5.28	0	1	0	0	0
13	APOB_HUMAN	(P04114) Apolipoprotein B-100 precursor (Apo B-100) [contains: apolipoprotein B-48 (Apo B-48)]	516666	6.61	19	4	0	0	0
14	APOC3_HUMAN	(P02656) Apolipoprotein C-III precursor (Apo-CIII) (ApoC-III)	10846	5.23	0	0	0	0	1
15	APOE_HUMAN	(P02649) Apolipoprotein E precursor (Apo-E)	36246	5.65	0	0	0	1[a]	1[a]
16	APOH_HUMAN	(P02749) Beta-2-glycoprotein I precursor (apolipoprotein H) (Apo-H) (B2GPI) (beta(2)GPI)	39584	8.34	4	3	0	0	0
17	BAZ2B_HUMAN	(Q9UIF8) Bromodomain adjacent to zinc finger domain 2B (hWALp4)	222285	6.00	12	1	0	0	0
18	BCL6B_HUMAN	(Q8N143) B-cell CLL/lymphoma 6 member B protein (Bcl6-associated zinc finger protein)	52669	9.23	0	0	1	0	0
19	BPAEB_HUMAN	(Q8WXK8) Bullous pemphigoid antigen 1; isoform 7 (bullous pemphigoid antigen) (BPA)	347740	4.76	23	0	0	1	0

Table 2 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	Number of peptides matched			
						Con A	SNA-I	UEA-I	PHA-L
20	BRSK2_HUMAN	(Q8IWQ3) BR serine/threonine-protein kinase 2 (EC 2.7.1.37) (serine/threonine-protein kinase 29)	82183	8.97	1	1	0	0	0
21	BSN_HUMAN	(Q9UPA5) Bassoon protein (zinc-finger protein 231)	418223	7.30	5	0	0	1	0
22	C4BP_HUMAN	(P04003) C4b-binding protein alpha chain precursor (C4bp) (proline-rich protein) (PRP)	69042	7.15	3	1	0	0	0
23	CAD12_HUMAN	(P55289) Brain-cadherin precursor (BR-cadherin) (cadherin-12) (*N*-cadherin 2)	88619	4.67	4	1[a]	1[a]	1[a]	0
24	CERU_HUMAN	(P00450) Ceruloplasmin precursor (EC 1.16.3.1) (ferroxidase)	122983	5.44	7	7	2	0	0
25	CFAB_HUMAN	(P00751) Complement factor B precursor (EC 3.4.21.47) (C3/C5 convertase) (properdin factor B)	86847	6.67	5	3	1	0	0
26	CFAH_HUMAN	(P08603) Complement factor H precursor (H factor 1)	143710	6.28	9	1	0	0	0
27	CLUS_HUMAN	(P10909) Clusterin precursor (complement-associated protein SP-40;40) (complement cytolysis inhibitor)	53031	5.89	6	1	0	0	0
28	CO3_HUMAN	(P01024) Complement C3 precursor [contains: C3a anaphylatoxin]	188585	6.02	3	16	1	3	2
29	CO3A1_HUMAN	(P02461) Collagen alpha 1(III) chain precursor	139724	6.18	4	1	0	0	0
30	CO4_HUMAN	(P01028) Complement C4 precursor [contains: C4a anaphylatoxin; C4b]	194247	6.66	4	3	1[b]	1[a]	1[a]
31	CO6_HUMAN	(P13671) Complement component C6 precursor	108425	6.31	2	1	0	0	0
32	CO8A_HUMAN	(P07357) Complement component C8 alpha chain precursor (complement component 8 alpha subunit)	66832	6.07	2	1	0	0	0
33	CP27B_HUMAN	(O15528) 25-hydroxyvitamin D-1 alpha hydroxylase; mitochondrial precursor (EC 1.14.13.13)	56982	9.34	1	1	0	0	0
34	DDEF1_HUMAN	(Q9ULH1) 130-kDa phosphatidylinositol 4,5-biphosphate-dependent ARF1 GTPase-activating protein	106815	6.80	2	0	1	0	0
35	DESP_HUMAN	(P15924) Desmoplakin (DP) (250/210 kDa paraneoplastic pemphigus antigen)	334023	6.44	20	1	0	0	0
36	DGKG_HUMAN	(P49619) Diacylglycerol kinase; gamma (EC 2.7.1.107) (diglyceride kinase) (DGK-gamma)	90535	6.29	2	0	0	1	0
37	DYH5_HUMAN	(Q8TE73) Ciliary dynein heavy chain 5 (axonemal beta dynein heavy chain 5) (HL1)	532488	5.79	23	0	0	1	0
38	EDG1_HUMAN	(P21453) Sphingosine 1-phosphate receptor Edg-1	43409	9.55	6	0	0	1	1

Table 2 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	Number of peptides matched			
						Con A	SNA-I	UEA-I	PHA-L
39	ENPL_HUMAN	(P14625) Endoplasmin precursor (94 kDa glucose-regulated protein) (GRP94) (gp96 homolog)	92696	4.76	6	0	0	1	0
40	ERCC4_HUMAN	(Q92889) DNA repair endonuclease XPF (EC 3.1.-.-) (DNA excision repair protein ERCC-4)	104136	6.42	2	0	0	1	0
41	FBX41_HUMAN	(Q8TF61) F-box only protein 41	95934	8.60	0	1	0	0	0
42	HEMO_HUMAN	(P02790) Hemopexin precursor (beta-1B-glycoprotein)	52385	6.55	5	5	5	0	1
43	HPT_HUMAN	(P00738) Haptoglobin precursor	45861	6.13	4	10	12	4	9
44	HPTR_HUMAN	(P00739) Haptoglobin-related protein precursor	39496	6.41	3	4	3	2	3
45	HRX_HUMAN	(Q03164) Zinc finger protein HRX (ALL-1) (trithorax-like protein)	436044	9.22	33	1[a]	0	0	1[a]
46	HV3S_HUMAN	(P01780) Ig heavy chain V-III region JON	12669	9.39	1	0	1	0	0
47	IC1_HUMAN	(P05155) Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh)	55347	6.09	6	4	1	0	0
48	IGHA1_HUMAN	(P01876) Ig alpha-1 chain C region	38486	6.08	2	8	7	6	5
49	IGHA2_HUMAN	(P01877) Ig alpha-2 chain C region	37283	5.71	4	2	2	0	0
50	IGHG1_HUMAN	(P01857) Ig gamma-1 chain C region	36596	8.46	1	6	9	7	7
51	IGHG2_HUMAN	(P01859) Ig gamma-2 chain C region	36489	7.66	1	1	4	3	1
52	IGHG3_HUMAN	(P01860) Ig gamma-3 chain C region (heavy chain disease protein) (HDC)	33222	7.89	3	0	1[a]	2	1[a]
53	IGHG4_HUMAN	(P01861) Ig gamma-4 chain C region	36431	7.18	1	0	0	2	0
54	IGJ_HUMAN	(P01591) Immunoglobulin J chain	16041	4.62	1	1	0	0	0
55	IRF7_HUMAN	(Q92985) Interferon regulatory factor 7 (IRF-7)	55042	5.69	0	0	0	1	0
56	ITIH1_HUMAN	(P19827) Inter-alpha-trypsin inhibitor heavy chain H1 precursor (ITI heavy chain H1)	101782	6.31	3	0	1	0	0
57	ITIH2_HUMAN	(P19823) Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy chain H2)	106826	6.40	3	0	0	0	1
58	ITIH4_HUMAN	(Q14624) Inter-alpha-trypsin inhibitor heavy chain H4 precursor (ITI heavy chain H4)	103522	6.51	4	0	1	0	0
59	K1C10_HUMAN	(P13645) Keratin; type I cytoskeletal 10 (cytokeratin 10) (K10) (CK 10)	59711	5.13	2	1	5	8	9
60	K1C9_HUMAN	(P35527) Keratin; type I cytoskeletal 9 (cytokeratin 9) (K9) (CK 9)	62178	5.14	3	0	2	6	3
61	K1CP_HUMAN	(P08779) Keratin; type I cytoskeletal 16 (cytokeratin 16) (K16) (CK 16)	51447	4.98	0	0	0	1[b]	2[a]

Table 2 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	Number of peptides matched			
						Con A	SNA-I	UEA-I	PHA-L
62	K22E_HUMAN	(P35908) Keratin; type II cytoskeletal 2 epidermal (cytokeratin 2e) (K2e) (CK 2e)	66111	8.07	2	0	2	3	4
63	K2C1_HUMAN	(P04264) Keratin; type II cytoskeletal 1 (cytokeratin 1) (K1) (CK 1) (67 kDa cytokeratin)	66018	8.16	3	2	7	13	12
64	K2C5_HUMAN	(P13647) Keratin; type II cytoskeletal 5 (cytokeratin 5) (K5) (CK 5) (58 kDa cytokeratin)	62637	8.14	3	0	0	0	1
65	K2C6A_HUMAN	(P02538) Keratin; type II cytoskeletal 6A (cytokeratin 6A) (CK 6A) (K6a keratin)	60162	8.14	2	0	0	0	3
66	K2C6C_HUMAN	(P48666) Keratin; type II cytoskeletal 6C (cytokeratin 6C) (CK 6C) (K6c keratin)	60317	8.14	2	0	0	1	0
67	KAC_HUMAN	(P01834) Ig kappa chain C region	11773	5.58	0	3	6	5	6
68	KLHL3_HUMAN	(Q9UH77) Kelch-like protein 3	65784	5.29	3	1	0	0	0
69	KNG_HUMAN	(P01042) Kininogen precursor (alpha-2-thiol proteinase inhibitor) [contains: Bradykinin]	72984	6.34	4	1[a]	2	0	0
70	KV2A_HUMAN	(P01614) Ig kappa chain V-II region Cum	12782	5.28	0	0	1	0	0
71	KV3A_HUMAN	(P01619) Ig kappa chain V-III region B6	11742	9.34	0	0	1	0	0
72	KV3B_HUMAN	(P01620) Ig kappa chain V-III region SIE	11882	8.70	0	1	0	2	1
73	LAC_HUMAN	(P01842) Ig lambda chain C regions	11401	6.91	0	4	4	2	4
74	LARGE_HUMAN	(O95461) Glycosyltransferase-like protein LARGE (EC 2.4.-.-)	88694	7.89	5	0	0	1	0
75	LRP2_HUMAN	(P98164) Low-density lipoprotein receptor-related protein 2 precursor (megalin) (glycoprotein 330)	540349	4.89	39	0	0	1	0
76	LV1I_HUMAN	(P06888) Ig lambda chain V-I region EPS	11521	9.39	1	0	0	1	0
77	LV3B_HUMAN	(P80748) Ig lambda chain V-III region LOI	12042	4.94	0	0	1	0	0
78	MAGD4_HUMAN	(Q96JG8) Melanoma-associated antigen D4 (MAGE-D4 antigen) (MAGE-E1 antigen)	81555	6.34	2	1	0	0	0
79	MARE3_HUMAN	(Q9UPY8) Microtubule-associated protein RP/EB family member 3 (end-binding protein 3) (EB3)	32247	5.33	4	1	0	0	0
80	MCM3A_HUMAN	(O60318) 80 kDa MCM3-associated protein (GANP protein)	220662	5.99	12	0	0	0	1
81	MED4_HUMAN	(Q9NPJ6) Mediator of RNA polymerase II transcription subunit 4 (Vitamin D3 receptor-interacting protein)	29784	5.02	2	0	1	0	0
82	MUC_HUMAN	(P01871) Ig mu chain C region	50210	6.35	5	2	0	0	1

Table 2 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	Number of peptides matched			
						Con A	SNA-I	UEA-I	PHA-L
83	MUCB_HUMAN	(P04220) Ig mu heavy chain disease protein (BOT)	43543	5.13	4	1[b]	1[a]	0	1[a]
84	MYOF_HUMAN	(Q9NZM1) Myoferlin (Fer-1 like protein 3)	236100	5.84	7	0	0	1	0
85	P85A_HUMAN	(P27986) Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase p85-alpha subunit)	83888	5.90	7	0	0	0	1
86	PCLO_HUMAN	(Q9Y6V0) Piccolo protein (aczonin)	568304	6.12	14	1	0	1	0
87	PDE8B_HUMAN	(O95263) High-affinity cAMP-specific and IBMX-insensitive 3′,5′-cyclic phosphodiesterase 8B	100456	6.35	4	0	1	0	0
88	PKDRE_HUMAN	(Q9NTG1) Polycystic kidney disease and receptor for egg jelly related protein precursor	257908	9.26	19	1	0	0	0
89	PLD2_HUMAN	(O14939) Phospholipase D2 (EC 3.1.4.4) (PLD 2) (choline phosphatase 2) (phosphatidylcholine-hydroly	106719	7.41	2	0	1	0	0
90	PRDM2_HUMAN	(Q13029) PR-domain zinc finger protein 2 (retinoblastoma protein-interacting zinc-finger protein)	191135	7.02	12	0	0	1	0
91	PZP_HUMAN	(P20742) Pregnancy zone protein precursor	165215	5.97	11	0	1	0	0
92	RASL2_HUMAN	(O43374) Ras GTPase-activating protein 4 (RasGAP-activating-like protein 2)	91427	8.01	3	0	1[a]	1[a]	1[a]
93	SACS_HUMAN	(Q9NZJ4) Sacsin	441658	6.81	25	1	1	1	1
94	SAFB2_HUMAN	(Q14151) Scaffold attachment factor B2	107921	5.84	3	0	0	0	1
95	SCG3_HUMAN	(Q8WXD2) Secretogranin-3 precursor (secretogranin III) (SgIII) (UNQ2502/PRO5990)	52973	4.94	3	0	0	1	0
96	SIP1_HUMAN	(O60315) Zinc finger homeobox protein 1b (smad interacting protein 1) (SMADIP1) (HRIHFB2411)	137844	5.90	10	1	0	0	0
97	THBG_HUMAN	(P05543) Thyroxine-binding globulin precursor (T4-binding globulin)	46637	5.87	4	1	0	0	0
98	THRB_HUMAN	(P00734) Prothrombin precursor (EC 3.4.21.5) (coagulation factor II)	71475	5.64	4	1	0	0	0
99	TLN2_HUMAN	(Q9Y4G6) Talin 2	273723	5.42	14	0	1	0	0
100	TRFE_HUMAN	(P02787) Serotransferrin precursor (transferrin) (siderophilin) (beta-1-metal binding globulin)	79280	6.81	2	11	13	10	8
101	TSP4_HUMAN	(P35443) Thrombospondin-4 precursor	108482	4.44	3	0	0	0	1
102	TTHY_HUMAN	(P02766) Transthyretin precursor (prealbumin) (TBPA) (TTR) (ATTR)	15991	5.52	1	0	0	1[b]	1[a]
103	U2AFM_HUMAN	(Q15696) U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit related-protein 2 (U2R)	58636	9.75	1	0	0	1	0

Table 2 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	Number of peptides matched			
						Con A	SNA-I	UEA-I	PHA-L
104	VIME_HUMAN	(P08670) Vimentin	53545	5.06	4	0	1	0	0
105	VTDB_HUMAN	(P02774) Vitamin D-binding protein precursor (DBP) (group-specific component) (Gc-globulin) (VDB)	54526	5.40	1	0	0	1	0
106	Y0552_HUMAN	(O60299) Hypothetical protein KIAA0552	72260	7.56	1	0	0	1	0
107	ZA2G_HUMAN	(P25311) Zinc-alpha-2-glycoprotein precursor (Zn-alpha-2-glycoprotein) (Zn-alpha-2-GP)	34079	5.57	4	0	0	0	1
108	ZN434_HUMAN	(Q9NX65) Zinc finger protein 434	56148	9.03	5	1	0	0	0

[a] Indicating the same peptide.
[b] Indicating different peptide.

glycosylated structures to glycosylated proteins; or (4) due to the fact that some more abundant proteins (e.g., albumin or its fragments) can feature glycation (non-enzymatic incorporation of glucose) that may result in lectin binding [32]. For example, several proteins, lacking the *N*-glycosylation motif were identified (Table 3) as a result of enrichment: apolipoproteins, and the immunoglobulin kappa and lambda chains. Their binding could possibly be explained by their *O*-glycosylation [33].

There has been a considerable debate in the proteomics and bioanalytical communities about the desirability of immunoaffinity depletion of blood serum. While a targeted removal of the major blood proteins seems distinctly beneficial in reducing the overall sample complexity and extending the dynamic range of searches for trace biomarkers, there is some risk of losing some proteins of interest due to their association with depleted proteins [23]. For example, serum albumin is known to act as a transport protein, associating with a wide variety of less abundant proteins. In simplifying the sample handling steps, no immunoaffinity depletion was performed in this work prior to the lectin affinity enrichment. As was proposed recently [10], the lectin affinity for glycoproteins, which represent a considerable fraction of targeted trace proteins, can be used as a means of extending the dynamic range. Accordingly, the following discussion will be based on comparison with the previously published data and the possible relation of protein localization to their abundance [34].

It is widely assumed that the proteins present in human blood serum at high concentration are predominantly, if not exclusively, extracellular proteins, while many low-abundance proteins are located inside the cells or in cellular membranes. The distribution of proteins identified in this study, in terms of their localization, is depicted in Fig. 1, where almost half of the identified proteins are of extracellular origin (41.58%), while 14.85% originate in cytoplasm. Some identified proteins are localized in nucleus (11.88%), plasma membrane (10.89%), endoplasmic reticulum (1.98%), mitochondrion (1.98%), Golgi apparatus (0.99%), cytoskeleton (1.98%), and secretory granule (0.99%). The location of the remaining 12.87% cannot be readily specified. The levels of some of the identified proteins are listed in Table 3 together with their carbohydrate contents. Carbohydrate content and the natural abundance of the glycoproteins summarized in Table 3 are listed in ref. [33]. It appears that the use of lectin enrichment permits the glycoprotein analysis within a concentration range spanning over 7–10 orders of magnitude. Even a wider concentration range is indicated by the identification of trace proteins, as suggested by their localization.

The molecular weight and p*I* distributions of proteins found in the lectin-bound fractions are illustrated in Fig. 2A and B, respectively. Molecular weights of the identified proteins span across a respectable range, as seen in Fig. 2A. However, the highest number of proteins was found with the molecular weight ranging from 10 to 100 kDa, while only a limited number of proteins with very high molecular weights (160–600 kDa) were identified. No proteins with a molecular weight greater than 600 kDa were observed in the lectin-bound fractions. As shown

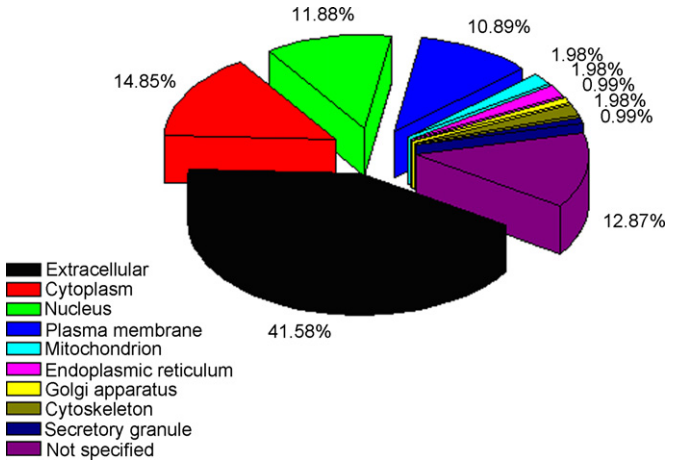

Fig. 1. Localization of proteins found in the bound fractions after high-performance lectin affinity enrichment of human blood serum using silica-based microcolumns loaded with four different lectins. Information was resourced from Human Protein Reference Database (www.hprd.org).

Table 3
Carbohydrate content and blood serum concentrations of some of the identified proteins listed in Tables 2 and 4

No.	Abbreviation	Description	Carbohydrate (%)[a]	Concentrations (pg/mL)[a]
1	ALBU_Human	Serum albumin precursor	0	4×10^{10}
α-Globulins				
2	A1AG1_Human	Alpha-1-acid glycoprotein 1 precursor (AGP 1) (orosomucoid 1)	42	3×10^{8}
3	A1AG2_Human	Alpha-1-acid glycoprotein 2 precursor (AGP 2) (orosomucoid 2)	42	6×10^{8}
4	A1AT_Human	Alpha-1-antitrypsin precursor (alpha-1-protease inhibitor) (alpha-1-antiproteinase)	13	3×10^{9}
5	A2MG_Human	Alpha-2-macroglobulin precursor (alpha-2-M)	9.4	1.4×10^{9}
6	AACT_Human	Alpha-1-antichemotrypsin precursor (ACT)	26	2×10^{8}
7	CERU_Human	Ceruloplasmin precursor (ferroxidase)	7	2.1×10^{8}
8	HPT_Human	Haptoglobin precursor	16.4–19.4	1.5×10^{9}
9	HPTR_Human	Heptoglobin-related protein precursor	N.D.	–
10	ITIH1_Human	Inter-alpha-trypsin inhibitor heavy chain H1 precursor (ITI heavy chain H1)	8	–
11	ITIH2_Human	Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy chain H2)	8	–
12	ITIH4_Human	Inter-alpha-trypsin inhibitor heavy chain H4 precursor (ITI heavy chain H4)	8	–
13	THBG_Human	Thyroxine-binding globulin precursor (T4-binding globulin)	20	$1.03–2.49 \times 10^{5}$
14	VTDB_Human	Vitamin D-binding protein precursor (DBP) (group-specific component) (Gc-globin) (VDB)	4	2×10^{3}
15	ZA2G_Human	Zinc-alpha-2-glycoprotein precursor (Zn-alpha-glycoprotein) (Zn-alpha-2-GP)	18	–
β-Globulin				
16	APOH_Human	Beta-2-glycoprotein I precursor (apoliprotein H)	19	1.1×10^{6}
17	HEMO_Human	Hemopexin precursor (beta-1B-glycoprotein)	22	1×10^{9}
18	TREF_Human	Serotransferrin precursor (siderophilin) (beta-1-metal binding globulin)	5.9	2.3×10^{9}
Coagulation proteins				
19	ANT3_Human	Antithrombin-III precursor (ATIII)	15	3.2×10^{8}
20	THRB_Human	Prothrombin precursor (coagulation factor II)	12	1.5×10^{8}
Apolipoproteins				
21	APOA1_Human	Apolipoprotein A-I precursor (Apo-AI)	*O*-glycosylated	1.4×10^{9}
22	APOA2_Human	Apolipoprotein A-II precursor (Apo-AII)	*O*-glycosylated	3×10^{8}
23	APOA4_Human	Apolipoprotein A-IV precursor (Apo-AIV)	*O*-glycosylated	–
24	APOB_Human	Apolipoprotein B-100 precursor (Apo-B-100)	*O*-glycosylated	$4–13 \times 10^{8}$
25	APOC3_Human	Apolipoprotein C-III precursor (Apo-CIII)	3 (*O*-glyco)	–
26	APOE_Human	Apolipoprotein E precursor (Apo-E)	1 (*O*-glyco)	3.4×10^{7}
27	CLUS_Human	Clusterin precursor (complement associate protein SP-40;40) (apolipoprotein J)	17–27	–
Immunoglobulin				
28	HV3S_Human	Ig heavy chain V-III region Jon	–	–
29	IGHA1_Human	Ig alpha-1 chain C region	8	$9.8–35.3 \times 10^{8}$
30	IGHA2_Human	Ig alpha-2 chain C region	8	$1.16–7.85 \times 10^{8}$
31	IGHG1_Human	Ig gamma-1 chain C region	3	$1.7–9.5 \times 10^{9}$
32	IGHG2_Human	Ig gamma-2 chain C region	3	$2.2–44 \times 10^{8}$
33	IGHG3_Human	Ig gamma-3 chain C region (heavy chain disease protein) (HDC)	3	$1.3–6.9 \times 10^{8}$
34	IGHG4_Human	Ig gamma-4 chain C region	3	$1–12 \times 10^{7}$
35	IGJ_Human	Immunoglobulin J chain	7.5	–
36	KAC_Human	Ig Kappa chain C region	*O*-glycosylated	$3.3–19.4 \times 10^{6}$
37	KV2A_Human	Ig Kappa chain V-II region Cum	*O*-glycosylated	–
38	KV3A_Human	Ig Kappa chain V-III region B6	*O*-glycosylated	–
39	KV3B_Human	Ig Kappa chain V-III region SIE	*O*-glycosylated	–
40	LAC-Human	Ig lambda chain C regions	*O*-glycosylated	$5.7–26.3 \times 10^{6}$
41	LV1I_Human	Ig lambda chain V-I regions EPS	*O*-glycosylated	–
42	LV3B_Human	Ig lambda chain V-III regions LOI	*O*-glycosylated	–
43	MUC-Human	Ig mu chain C region	10	1.3×10^{9}
44	MUCB_Human	Ig mu heavy chain disease protein (BOT)	10	1.3×10^{9}
Complement components and factors				
45	C4BP_Human	C4b-binding protein alpha chain precursor (C4bp) (prolin-rich protein) (PRP)	+	$2.46–5.1 \times 10^{8}$
46	CFAB_Human	Complement factor B precursor (C3/C5 convertase) (properdin factor B)	10	$1.7–4.2 \times 10^{8}$
47	CFAH_Human	Complement factor H (H factor 1)	9.3	$4.22–6.91 \times 10^{8}$
48	CO3_Human	Complement factor C3 precursor (contains: C3a anaphylatoxin)	2	$8.2–17 \times 10^{8}$
49	CO3A1_Human	Collagen alpha 1(III) chain precursor	2	–
50	CO4_Human	Complement C4 precursor (contains: C4a anaphylatoxin; C4b)	2	1.7×10^{8}
51	CO6_Human	Complement component C6 precursor	3.8	$6–13.5 \times 10^{7}$
52	CO8A_Human	Complement component C8 alpha chain precursor	2	$8.5–19.3 \times 10^{7}$

[a] Values obtained from ref. [33].

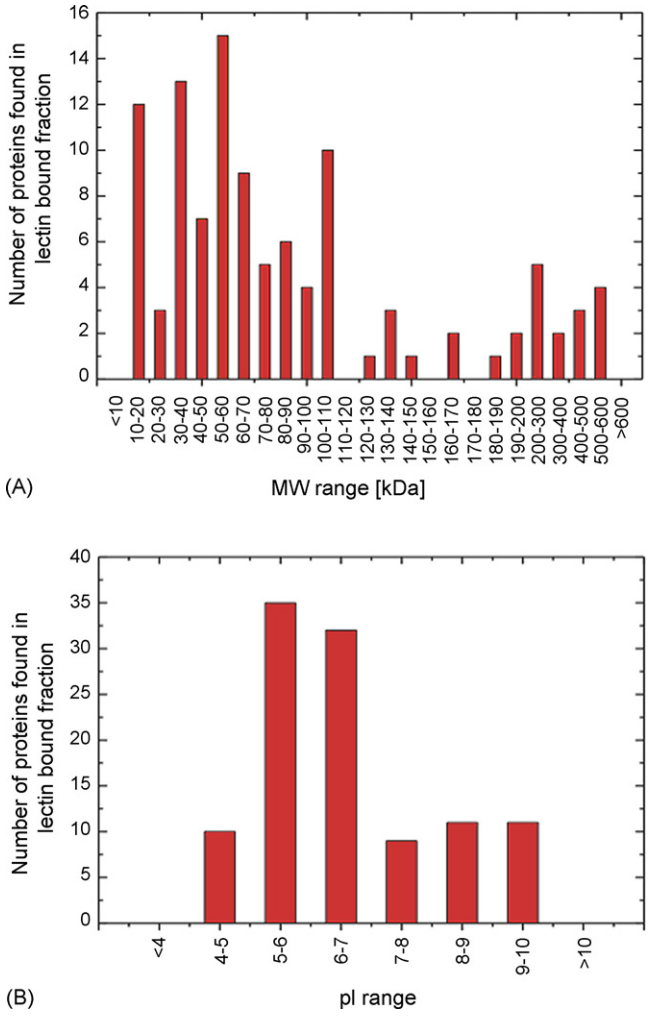

Fig. 2. Molecular weight (A) and p*I* distribution (B) of the proteins identified in the bound fractions after high-performance affinity enrichment of human blood serum on individual lectins at microscale.

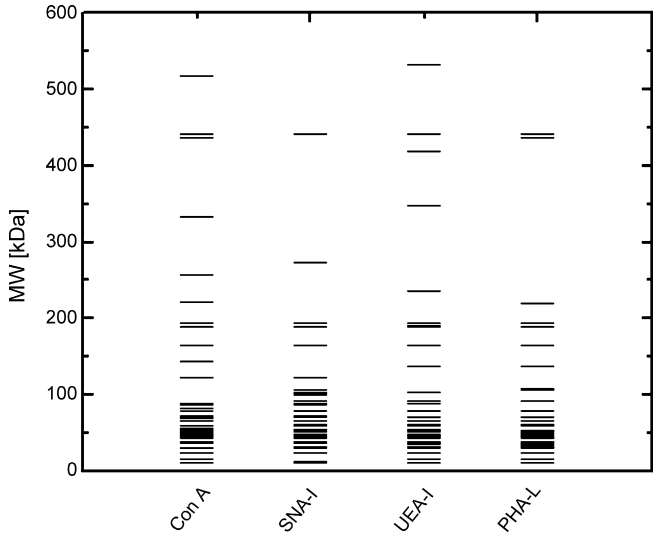

Fig. 3. Glycoproteomic map based on the molecular weight distribution of human blood serum proteins identified in the bound fractions after high-performance lectin affinity enrichment as a result of the different affinity towards the used lectins.

in Fig. 2B, no proteins with p*I*-values less than 4 or higher than 10 were identified. The p*I*-values of 27 identified proteins (>33%) ranged from 5 to 6, while 25 proteins (>31%) had p*I*-values higher than 6 but less than 7. These data are very similar to the results previously reported in combination with an ion-exchange chromatographic fractionation [35].

Using sequential lectin affinity chromatography with four different lectins has allowed us to create a distribution map of various glycoproteins in the individual bound fractions. Such a map, describing the differences in the affinity of glycoproteins toward individual lectins, is depicted in Fig. 3. Here, the proteins showing affinity to one specific lectin are represented as lines forming the column. A substantial preference of some proteins within a certain molecular weight range will thus be reflected in the areas with a higher line density. There were 55 proteins which demonstrated affinity to Con A and 47 proteins that were bound to SNA-I. Additionally, 49 proteins were recognized by UEA-I lectin, while 46 proteins were bound to the PHA-L lectin. These numbers reflect an extensive overlap among the individual lectin fractions. This is likely to originate from the heterogeneity of glycan structures associated with particular glycoproteins, so that an interaction of certain glycoproteins with more than one lectin is highly probable.

As indicated by our data, the variation of protein affinity toward the selected lectins followed the expected trend resulting from the difference in lectin specificities. The majority of glycoproteins were bound to Con A lectin, which is known to feature the broadest specificity. The other lectins, possessing a more restricted specificity, resulted in a lower number of bound proteins. A molecular weight distribution of glycoproteins also changed with respect to the type of a binding lectin. This variation is apparent from the densities of the lines representing the molecular weights of identified proteins. According to the map depicted in Fig. 3, all lectins seem to prefer binding glycoproteins with molecular weights less than 100 kDa: Con A preferred glycoproteins with the molecular masses within 50–100 kDa range, SNA-I and UEA-I binding profiles were spread across the entire 10–100 kDa range, while PHA-L predominantly recognized proteins around 50 kDa. This being said, the molecular weights depicted in Fig. 3 are that of protein backbone. Lectin enrichment is independent of protein molecular weight, but rather influenced by the nature of the glycan moieties and their extent.

As demonstrated previously by Hancock and co-workers [20], treating glycoproteins with specific enzymes (such as sialidases and fucosidases) can modulate a glycoprotein profile of human blood serum, resulting in the change of affinity to fucose- and sialic-acid-specific lectins. Such a phenomenon would probably appear as a shift in molecular weight distribution of a lectin-bound fraction, so that some lines due to sialylated or fucosylated glycoproteins would either show up in a different column, or totally disappear. Whether this strategy would enhance or hinder the diagnostic information remains to be elucidated. The alteration of a glycan heterogeneity in a glycoprotein in blood serum, which could result as a consequence of diseases, such as cancer, could be monitored through similar

Table 4
List of proteins identified in the bound fraction after high-performance multi-lectin affinity chromatography of human blood serum

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	No. of peptides matched
1	A1AG1_HUMAN	(P02763) Alpha-1-acid glycoprotein 1 precursor (AGP 1) (orosomucoid 1) (OMD 1)	23725	4.93	5	2
2	A1AG2_HUMAN	(P19652) Alpha-1-acid glycoprotein 2 precursor (AGP 2) (orosomucoid 2) (OMD 2)	23873	5.03	5	1
3	A1AT_HUMAN	(P01009) Alpha-1-antitrypsin precursor (alpha-1 protease inhibitor) (alpha-1-antiproteinase)	46878	5.37	3	4
4	A2MG_HUMAN	(P01023) Alpha-2-macroglobulin precursor (alpha-2-M)	164600	6.00	8	31
5	AACT_HUMAN	(P01011) Alpha-1-antichymotrypsin precursor (ACT)	47792	5.33	6	4
6	ADSV_HUMAN	(Q9Y6U3) Adseverin (scinderin)	80895	5.5	4	1
7	ALBU_HUMAN	(P02768) Serum albumin precursor	71317	5.92	0	34
8	ANGT_HUMAN	(P01019) Angiotensinogen precursor [contains: angiotensin I (Ang I); angiotensin II (Ang II)	53406	5.87	4	1
9	AP4B1_HUMAN	(Q9Y6B7) Adapter-related protein complex 4 beta 1 subunit (beta subunit of AP-4)	84576	5.59	0	1
10	APOA1_HUMAN	(P02647) Apolipoprotein A-I precursor (Apo-AI) (ApoA-I)	30759	5.56	0	6
11	APOA2_HUMAN	(P02652) Apolipoprotein A-II precursor (Apo-AII) (ApoA-II)	11282	6.27	0	2
12	APOA4_HUMAN	(P06727) Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)	45343	5.28	0	1
13	APOB_HUMAN	(P04114) Apolipoprotein B-100 precursor (Apo B-100) [contains: apolipoprotein B-48 (Apo B-48)]	516666	6.61	19	1
14	APOH_HUMAN	(P02749) Beta-2-glycoprotein I precursor (apolipoprotein H) (Apo-H) (B2GPI) (beta(2)GPI)	39584	8.34	4	1
15	ARIA10_HUMAN	(P29374) Retinoblastoma-binding protein 1 (RBBP-1)	143548	5.01	10	1
16	CERU_HUMAN	(P00450) Ceruloplasmin precursor (EC 1.16.3.1) (ferroxidase)	122983	5.44	7	1
17	CFAB_HUMAN	(P00751) Complement factor B precursor (EC 3.4.21.47) (C3/C5 convertase) (properdin factor B)	86847	6.67	5	1
18	CFAH_HUMAN	(P08603) Complement factor H precursor (H factor 1)	143710	6.28	9	4
19	CO3_HUMAN	(P01024) Complement C3 precursor [Contains: C3a anaphylatoxin]	188585	6.02	3	36
20	CO4_HUMAN	(P01028) Complement C4 precursor [contains: C4a anaphylatoxin; C4b]	194247	6.66	4	8

Table 4 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	No. of peptides matched
21	CO5_HUMAN	(P01031) Complement C5 precursor [contains: C5a anaphylatoxin]	189923	6.11	4	1
22	CO6_HUMAN	(P13671) Complement component C6 precursor	108425	6.31	2	1
23	CO8A_HUMAN	(P07357) Complement component C8 alpha chain precursor (complement component 8 alpha subunit)	66832	6.07	2	2
24	CO8G_HUMAN	(P07360) Complement component C8 gamma chain precursor	22377	8.86	0	1
25	CPN2_HUMAN	(P22792) Carboxypeptidase N 83 kDa chain precursor (carboxypeptidase N regulatory subunit)	61431	5.63	8	1
26	ECEL1_HUMAN	(O95672) Endothelin-converting enzyme-like 1 (EC 3.4.24.-) (Xce protein) (UNQ2431/PRO4991)	88560	6.56	3	1
27	EGFL4_HUMAN	(Q7Z7M0) Multiple EGF-like-domain protein 4 (multiple epidermal growth factor-like domains 8)	264669	6.53	27	1
28	FETUA_HUMAN	(P02765) Alpha-2-HS-glycoprotein precursor (fetuin-A) (alpha-2-Z-globulin) (Ba-alpha-2-glycoprotein)	40098	5.43	2	1
29	HEMO_HUMAN	(P02790) Hemopexin precursor (beta-1B-glycoprotein)	52385	6.55	5	7
30	HEP2_HUMAN	(P05546) Heparin cofactor II precursor (HC-II) (protease inhibitor leuserpin 2) (HLS2)	57205	6.41	3	1
31	HPT_HUMAN	(P00738) Haptoglobin precursor	45861	6.13	4	13
32	HPTR_HUMAN	(P00739) Haptoglobin-related protein precursor	39496	6.41	3	5
33	IC1_HUMAN	(P05155) Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh)	55347	6.09	6	3
34	IGHA1_HUMAN	(P01876) Ig alpha-1 chain C region	38486	6.08	2	8
35	IGHA2_HUMAN	(P01877) Ig alpha-2 chain C region	37283	5.71	4	2
36	IGHG1_HUMAN	(P01857) Ig gamma-1 chain C region	36596	8.46	1	9
37	IGHG2_HUMAN	(P01859) Ig gamma-2 chain C region	36489	7.66	1	3
38	IGHG3_HUMAN	(P01860) Ig gamma-3 chain C region (Heavy chain disease protein) (HDC)	33222	7.89	3	1
39	IGHG4_HUMAN	(P01861) Ig gamma-4 chain C region	36431	7.18	1	3
40	ITA5_HUMAN	(P08648) Integrin alpha-5 precursor (fibronectin receptor alpha subunit) (integrin alpha-F) (VLA-5)	115605	5.5	14	1
41	ITIH2_HUMAN	(P19823) Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy chain H2)	106826	6.40	3	1

Table 4 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	No. of peptides matched
42	ITIH4_HUMAN	(Q14624) Inter-alpha-trypsin inhibitor heavy chain H4 precursor (ITI heavy chain H4)	103522	6.51	4	1
43	K1C10_HUMAN	(P13645) Keratin; type I cytoskeletal 10 (cytokeratin 10) (K10) (CK 10)	59711	5.13	2	2
44	K1C9_HUMAN	(P35527) Keratin; type I cytoskeletal 9 (cytokeratin 9) (K9) (CK 9)	62178	5.19	3	1
45	K2C1_HUMAN	(P04264) Keratin; type II cytoskeletal 1 (cytokeratin 1) (K1) (CK 1) (67 kDa cytokeratin)	66018	8.16	3	1
46	K2C5_HUMAN	(P13647) Keratin; type II cytoskeletal 5 (cytokeratin 5) (K5) (CK 5) (58 kDa cytokeratin)	62637	8.14	3	1
47	KAC_HUMAN	(P01834) Ig kappa chain C region	11773	5.58	0	4
48	KIF1A_HUMAN	(Q12756) Kinesin-like protein KIF1A (axonal transporter of synaptic vesicles)	192561	5.93	9	1
49	KNTC1_HUMAN	(P50748) Kinetochore-associated protein 1 (rough deal homolog) (hRod) (HsROD) (Rod)	253211	5.67	7	1
50	KV3B_HUMAN	(P01620) Ig kappa chain V-III region SIE	11882	8.70	0	1
51	LAC_HUMAN	(P01842) Ig lambda chain C regions	11401	6.91	0	4
52	MCM3_HUMAN	(P25205) DNA replication licensing factor MCM3 (DNA polymerase alpha holoenzyme-associated protein	91551	5.53	2	1
53	MUC_HUMAN	(P01871) Ig mu chain C region	50210	6.35	5	1
54	MUCB_HUMAN	(P04220) Ig mu heavy chain disease protein (BOT)	43543	5.13	4	1
55	NOLC1_HUMAN	(Q14978) Nucleolar phosphoprotein p130 (nucleolar 130 kDa protein) (140 kDa nucleolar phosphoprotei)	73677	9.48	4	1
56	PRRP_HUMAN	(P81277) Prolactin-releasing peptide precursor (PrRP) (prolactin-releasing hormone)	9747	11.66	0	1
57	PTN1_HUMAN	(P18031) Tyrosine-protein phosphatase; non-receptor type 1 (EC 3.1.3.48)	50505	5.88	2	1
58	RASL2_HUMAN	(O43374) Ras GTPase-activating protein 4 (RasGAP-activating-like protein 2)	91427	8.01	3	1
59	SACS_HUMAN	(Q9NZJ4) Sacsin	441658	6.81	25	1
60	SAMP_HUMAN	(P02743) Serum amyloid P-component precursor (SAP) (9.5S alpha-1-glycoprotein)	25485	6.10	2	1
61	SEPT1_HUMAN	(Q8WYJ6) Septin 1 (LARP) (serologically defined breast cancer antigen NY-BR-24)	42400	5.56	0	1

Table 4 (*Continued*)

No.	Abbreviation	Description	Mass	p*I*	No. of *N*-glycosylation motifs	No. of peptides matched
62	THBG_HUMAN	(P05543) Thyroxine-binding globulin precursor (T4-binding globulin)	46637	5.87	4	1
63	TRFE_HUMAN	(P02787) Serotransferrin precursor (transferrin) (siderophilin) (beta-1-metal binding globulin)	79280	6.81	2	6
64	TTC3_HUMAN	(P53804) Tetratricopeptide repeat protein 3 (TPR repeat protein 3) (TPR repeat protein D)	232882	7.63	10	1
65	VTDB_HUMAN	(P02774) Vitamin D-binding protein precursor (DBP) (group-specific component) (Gc-globulin) (VDB)	54526	5.40	1	1
66	ZFP91_HUMAN	(Q96JP5) Zinc finger protein 91 homolog (Zfp-91) (FKSG11 protein)	64261	7.02	1	1
67	ZN169_HUMAN	(Q14929) Zinc finger protein 169	68757	9.34	1	1

plots. Accordingly, such glycoproteomic maps could eventually become a valuable tool for a quick diagnosis of health disorders at their early stages.

3.1. Fractionation of human blood serum on multi-lectin affinity microcolumn

An experimental arrangement, referred to recently as "multi-lectin affinity chromatography" by Hancock and co-workers [19], is a selective enrichment mode in which the stationary phase is comprised of a mixture of different lectin-agarose beads. The idea behind this enrichment strategy is to use a set of lectins (in equal proportions) with different specificities, in one column, and trap most, if not all, glycoproteins from a complex glycosylated sample. The glycoproteins are later displaced from the column with a mixture of certain sugar haptens. Presumably, this enrichment mode should provide comparable results to those achieved here with the sequential use of different lectin columns, even though the possible advantage of a selective enrichment and fraction profiling is lost when all bound glycoproteins are simultaneously eluted from the preconcentration medium.

Using proportionally the same lectin composition in a single column (mimicking the multi-lectin experiment [19]), we have compared the enrichment capability of this mode with the sequential arrangement described above. Accordingly, the lectin-silica slurry was packed inside a 1 mm × 15-cm PEEK tubing, after which the same sample aliquot (20-μL serum volume) was injected and the retained proteins eluted. The proteins identified through the multi-lectin mode are listed in Table 4.

Surprisingly, fewer proteins were identified here (67 proteins, Table 4), when compared with the sequential-lectin arrangement (108 proteins, Table 2), indicating a lower trapping efficiency. Apparently, unlike the sequential enrichment on the individual microcolumns where the sample components are exposed to one lectin at a time, the protein access to the multiple-lectin beads might be more difficult due to a higher initial concentration when using the multi-lectin approach. Additionally, the substantial difference between the volume of elution buffer used in the sequential lectin approach, relative to the multi-lectin mode, may effectively reduce sample losses, thus enhancing sensitivity and detection thresholds for protein identification.

Only 11 proteins, of the 67 identified proteins, did not feature the *N*-glycosylation motif. The localization of these proteins is depicted in Fig. 4. The majority of identified proteins are extracellular (59%), which we consider relatively high in abundance, while 20% are localized equally in cytoplasm and nucleus. The proteins due to the plasma membrane amount to 7.6%. No proteins seem to be originating from the endoplasmic reticulum, Golgi apparatus and mitochondria, with the rest (13.6%) unspecified. A distribution of proteins identified through the

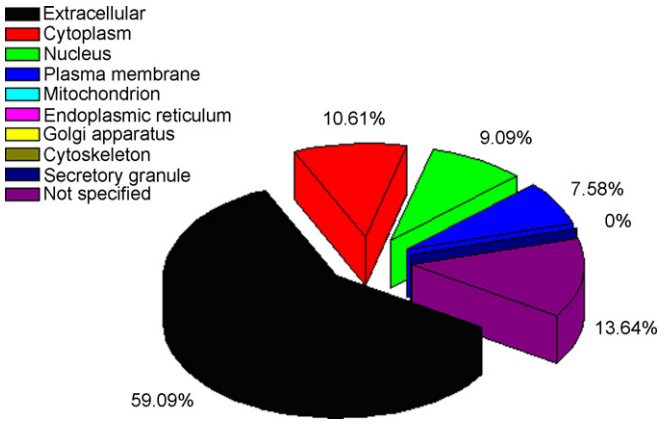

Fig. 4. Localization of proteins found in the bound fraction after high-performance multi-lectin affinity enrichment of human blood serum using a silica-based microcolumn. Information was resourced from Human Protein Reference Database (www.hprd.org).

multi-lectin approach in terms of mass range and the p*I* range was similar to that observed with the sequential-lectin approach (data not shown).

4. Conclusions

We have demonstrated here a selective enrichment of glycoproteins from microliter volumes of blood serum, employing lectin high-performance affinity chromatography on silica-based microcolumns. With eliminating a need for having large volumes of various biological fluids available for an experiment, our approach also substantially reduced the sample handling and, thereby, possible sample losses. We have also shown two alternatives of how lectin affinity enrichment of glycoproteins from real samples can be conducted at the microscale level. Knowing the differences in specificities of used lectins, this methodology can obviously be used for small-scale profiling of glycoproteins present in blood serum for diagnostic purposes. Moreover, high-performance lectin affinity chromatography utilized in a silica-based microcolumn format also offers the potential of being used as a part of valve-based high-pressure system, thus permitting on-line connection to high-resolution analytical techniques [17]. Development of an automated system is currently being investigated [36]. Moreover, the use of more advanced mass analyzers, such as for example LTQ-FT instrument, should at least double the number of identified proteins, as indicated recently [37].

Acknowledgements

This work was supported by Grant No. GM24349 from the National Institute of General Medical Sciences, U.S. Department of Health and Human Services and a center grant from the Indiana 21st Century Research and Technology Fund. The mass spectrometer used in this study was acquired as a result of support from Indiana Genomics Initiative (INGEN), which is funded in part by the Lilly Endowment Inc. This work was also supported by the National Institute of Health through the National Center for Research Resources (NCRR) by Grant Number RR018942 for the National Center for Glycomics and Glycoproteomics.

References

[1] S. Fukui, T. Feizi, C. Galustin, A.M. Lawoson, W. Chai, Nat. Biotechnol. 20 (2002) 1011.
[2] R.G. Gudepu, F. Wold, Proteins: Analysis and Design, Academic, San Diego, 1998, 121.
[3] P. van den Steen, P.M. Rudd, R.A. Dwek, G. Opdenakker, Crit. Rev. Biochem. Mol. Biol. 33 (1998) 151.
[4] J.W. Dennis, M. Granovsky, C.E. Warren, Bioassays 21 (1999) 412.
[5] J.B. Lowe, J.D. Marth, Ann. Rev. Biochem. 72 (2003) 643.
[6] R. Hjorth, Vretblad, in: R. Epton (Ed.), Lect. Chem. Soc. Int. Symp., Ellis Horwood Ltd., Uppsala, Sweden, 1976, p. 330.
[7] D.S. Hage, Clin. Chem. 45 (1999) 593.
[8] J. Nawarak, S. Phutrakul, S.-T. Chen, J. Proteome Res. 3 (2004) 383.
[9] L. Xiong, D. Andrews, F. Regnier, J. Proteome Res. 2 (2003) 618.
[10] Z. Yang, W.S. Hancock, T.R. Chew, L. Bonilla, Proteomics 5 (2005) 3353.
[11] R.D. Cummings, Methods Enzymol. 230 (1994) 66.
[12] X. Mao, Y. Luo, Z. Dai, K. Wang, Y. Du, B. Lin, Anal. Chem. 76 (2004) 6941.
[13] L. Hermo, R. Winikoff, F.W.K. Kan, Histochemistry 98 (1992) 93.
[14] J.L. Bundy, C. Fenselau, Anal. Chem. 73 (2001) 751.
[15] K.T. Pilobello, L. Krishnamoorthy, D. Slawek, L.K. Mahal, Chem. Biochem. 6 (2005) 985.
[16] M. Bedair, Z. El Rassi, J. Chromatogr. A 1079 (2005) 236.
[17] M. Madera, Y. Mechref, M.V. Novotny, Anal. Chem. 77 (2005) 4081.
[18] B. Zhang, M.M. Palcic, H. Mo, I.J. Goldstein, O. Hindsgaul, Glycobiology 11 (2001) 141.
[19] Z. Yang, W.S. Hancock, J. Chromatogr. A 1053 (2004) 79.
[20] Z. Yang, W.S. Hancock, J. Chromatogr. A 1070 (2005) 57.
[21] P.-O. Larsson, Methods Enzymol. 104 (1984) 212.
[22] M.M. Bradford, Anal. Biochem. 72 (1972) 248.
[23] A.K. Yocum, K. Yu, T. Oe, I.A. Blair, J. Proteome Res. 4 (2005) 1722.
[24] K.A. Resing, K. Meyer-Arendt, A.M. Mendoza, L.D. Aveline-Wolf, K.R. Jonscher, K.G. Pierce, W.M. Old, H.T. Cheung, S. Russell, J.L. Wattawa, G.R. Goehle, R.D. Knight, N.G. Ahn, Anal. Chem. 76 (2004) 3556.
[25] J. Turková, Bioaffinity Chromatography, Elsevier Science Publishers B.V., Amsterdam, 1993, p. 800.
[26] J.U. Baenziger, D. Fiete, J. Biol. Chem. 254 (1979) 2400.
[27] N. Shibuya, I.J. Goldstein, W.F. Broekaert, M. Nsimba-Lubaki, B. Peeters, W.J. Peumans, J. Biol. Chem. 262 (1987) 1596.
[28] S. Sugii, E.A. Kabat, H.H. Baer, Carbohydr. Res. 99 (1982) 99.
[29] R.D. Cummings, S. Kornfeld, J. Biol. Chem. 257 (1982) 11235.
[30] M.V. Dwek, H.A. Ross, A.J.C. Leathem, Proteomics 1 (2001) 756.
[31] R. Qiu, F.E. Regnier, Anal. Chem. 77 (2005) 2802.
[32] M.P. Cohen, S. Chen, F.N. Ziyadeh, E. Shea, E.A. Hud, G.T. Lautenslager, C.W. Shearman, Kidney Int. 68 (2005) 1554.
[33] J.U. Baenziger, in: F.W. Putnam (Ed.), The Plasma Proteins. Structure, Function, and Genetic Control, Academic Press, INC, Orlando, 1984, p. 271.
[34] L.N. Anderson, N.G. Anderson, Mol. Cell. Proteomics 1 (2002) 845.
[35] W.-H. Jin, J. Dai, S.-J. Li, Q.-C. Xia, H.-F. Zou, R. Zeng, J. Proteome Res. 4 (2005) 613.
[36] M. Madera, Y. Mechref, I. Klouckova, M.V. Novotny, J. Proteome Res. 5 (2006) 2348.
[37] J.S. Andersen, Y.W. Lam, A.K.L. Leung, S.-E. Ong, C.E. Lyon, A.I. Lamond, M. Mann, Nature 433 (2005) 77.