Glycoconj J (2012) 29:411–424
DOI 10.1007/s10719-012-9420-3

Surface glycoproteomic analysis of hepatocellular carcinoma cells by affinity enrichment and mass spectrometric identification

Wei Mi · Wei Jia · Zhaobin Zheng · Jinglan Wang · Yun Cai · Wantao Ying · Xiaohong Qian

Received: 14 April 2012 / Revised: 5 June 2012 / Accepted: 12 June 2012 / Published online: 1 July 2012

Abstract Cell surface glycoproteins are one of the most frequently observed phenomena correlated with malignant growth. Hepatocellular carcinoma (HCC) is one of the most malignant tumors in the world. The majority of hepatocellular carcinoma cell surface proteins are modified by glycosylation in the process of tumor invasion and metastasis. Therefore, characterization of cell surface glycoproteins can provide important information for diagnosis and treatment of liver cancer, and also represent a promising source of potential diagnostic biomarkers and therapeutic targets for hepatocellular carcinoma. However, cell surface glycoproteins of HCC have been seldom identified by proteomics approaches because of their hydrophobic nature, poor solubility, and low abundance. The recently developed cell surface-capturing (CSC) technique was an approach specifically targeted at membrane glycoproteins involving the affinity capture of membrane glycoproteins using glycan biotinylation labeling on intact cell surfaces. To characterize the cell surface glycoproteome and probe the mechanism of tumor invasion and metastasis of HCC, we have modified and evaluated the cell surface-capturing strategy, and applied it for surface glycoproteomic analysis of hepatocellular carcinoma cells. In total, 119 glycosylation sites on 116 unique glycopeptides were identified, corresponding to 79 different protein species. Of these, 65 (54.6 %) new predicted glycosylation sites were identified that had not previously been determined experimentally. Among the identified glycoproteins, 82 % were classified as membrane proteins by a database search, 68 % had transmembrane domains (TMDs), and 24 % were predicted to contain 2–13 TMDs. Moreover, a total of 26 CD antigens with 50 glycopeptides were detected in the membrane glycoproteins of hepatocellular carcinoma cells, comprising 43 % of the total glycopeptides identified. Many of these identified glycoproteins are associated with cancer such as CD44, CD147 and EGFR. This is a systematic characterization of cell surface glycoproteins of HCC. The membrane glycoproteins identified in this study provide very useful information for probing the mechanism of liver cancer invasion and metastasis.

Keywords Surface glycoproteins · Hepatocellular carcinoma cells · Mass spectrometry

Abbreviations

MS	mass spectrometry
MS/MS	tandem mass spectrometry
LC	liquid chromatography
peptide-N-glycosidase F	PNGase F
TMDs	transmembrane domains
GRAVY	grand average of hydrophobicity

Wei Mi and Wei Jia The authors contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10719-012-9420-3) contains supplementary material, which is available to authorized users.

W. Mi · W. Jia · Z. Zheng · J. Wang · Y. Cai · W. Ying (✉) · X. Qian (✉)
State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine,
33 Life Park Road, Changping District, Beijing 102206, People's Republic of China
e-mail: yingwtll@yahoo.com.cn
e-mail: qianxh1@163.com

W. Mi
Division of Medical and Biological Measurements, National Institute of Metrology,
No.18, Bei San Huan Dong Lu,
Beijing 100013, People's Republic of China

W. Jia
Institute of Biophysics, Chinese Academy of Science,
No. 15 Datun Road, Chaoyang District, Beijing 100101, China

Introduction

Cell surface membrane proteins straddle the lipid bilayer boundary between the cell and the environment surrounding it, and they play many important roles in fundamental biological processes. Membrane proteins have been extensively targeted for drug design, and they account for 70 % of all known drug targets [1–3]. One feature of mammalian membrane proteins is that most of them are extensively glycosylated [4, 5]. Glycosylation of membrane proteins plays a key role in biological processes such as signal transduction, cell adhesion and cell recognition. Some membrane glycoproteins have been reported to be involved in cancer progression and immune responses [6–8] in carcinomas [9]. Hepatocellular carcinoma (HCC) is one of the most malignant tumors in the world due to high rate of metastasis and relapse, low rate of 5-year patient survival and poor prognosis. Including hepatocellular carcinoma cells, the majority of the tumor cell membrane proteins are modified by glycosylation, but also in the process of tumor invasion and metastasis with the change of disease process. The adhesion and receptor factors associated with HCC metastasis are mostly membrane glycoproteins [10]. Therefore, proteomic analysis of cell surface glycoproteins for hepatocellular carcinoma cells can provide important information for the early diagnosis of HCC, process monitoring, prognostic assessment and exploring of the therapeutic targets.

Many researchers have focused on distinct subsets of membrane glycoproteins. Unfortunately, the purification of membrane glycoproteins is particularly difficult. Their low solubility and low abundance has severely limited the proteomic profiling of this class of proteins from complex biological mixtures. The advances in membrane proteomics [11] and glycoproteomics [12] have enabled the study of membrane glycoproteins from a variety of membrane preparations with impressive results. Several species of lectins and hydrazine have been used in the second step of the enrichment of cell surface glycoproteins following the initial step of cell membrane purification by ultracentrifugation [13, 14]. As membranes often lose their specific structure upon cell lysis, a typical membrane-rich fraction prepared by ultracentrifugation is heavily contaminated with other membrane components from subcellular compartments [15]. The reliability of the identification results obtained using this strategy is primarily dependent on the quality of the membrane fraction. Furthermore, the purification procedures for cell surface membranes are complicated and time-consuming. Thus, we questioned whether alternative purification strategies could be developed that did not rely on membrane purification. As membrane glycoproteins are positioned in the cell membrane with their glycans on the outside of the cell, they are naturally separated from other components of the cell. Thus, the position of these glycans offers the opportunity for direct labeling and enrichment of membrane glycoproteins on the cell surface [16, 17].

In this paper, we modified and applied a highly specific and simple approach for surface glycoproteomic analysis of the hepatocellular carcinoma cells by targeted enrichment and identification of cell membrane glycoproteins. The method involved several steps. First, the cell surface membrane glycoproteins of intact cells were mildly oxidized and specifically labeled with a biotin group. Second, after cell lysis, a simple centrifugation process, protein extraction and tryptic digestion, the biotinylated membrane glycopeptides were enriched via affinity capture using immobilized avidin. Finally, the samples were separated and identified by LC-MS. The feasibility of the biotin label and capture was demonstrated by standard glycoproteins. We applied the strategy to surface glycoproteomic analysis of the hepatocellular carcinoma cells. In total, 119 glycosylation sites from 116 glycopeptides were identified, corresponding to 79 different protein species. Over half (65) of the N-glycosylation sites identified had not been experimentally validated in the Uniprot database, with 49 potential glycosylation sites and 16 unknown sites. Many cell surface glycoproteins functionally associated with hepatocellular carcinoma such as CD44, CD147 were identified. These glycoproteins on the cell surface play their roles associated with cell adhesion, proliferation, and motility. They are correlated with tumor prognosis and regulate tumor progression and metastasis. The strategy and the data were expected to facilitate the researches on membrane glycoproteome and biomarker discovery for hepatocellular carcinoma.

Materials and methods

Chemicals and reagents

All standard proteins and periodate were obtained from Sigma-Aldrich (St. Louis, MO). Tris(2-carboxyethyl)phosphine (TCEP) and EZ-Link Biotin-LC-Hydrazide were obtained from Pierce (Rockford, US). Protease inhibitor cocktail and peptide-N-glycosidase F (PNGase F) were from Roche (Mannheim, Germany). Cell culture reagents and media were from Hyclone (Logan, UT). Sequence grade trypsin was from Promega (Madison, WI). All other chemicals were purchased from Thermo Fisher Scientific (Waltham, MA).

Cell culture and cell surface labeling

The hepatocellular carcinoma cell line HepG2 was grown at 37 °C in Dulbecco's modified Eagle's medium (DMEM) with 10 % FBS until approaching confluency (~90 %). Then, 1×10^8 cells were washed with PBS (10 mM

NaH_2PO_4/Na_2HPO_4, pH 7.2, 138 mM NaCl, 2.7 mM KCl) three times, and then incubated in the dark at 4 °C for 30 min with a solution of 10 mM sodium periodate in cold PBS (pH 7.2). The reaction solution was removed and the cells were washed with PBS three times. Then, the cells were incubated with 200 μg/mL EZ-Link Biotin-LC-Hydrazide solution in PBS (pH 7.2) for 1 h at room temperature with gentle shaking. After removal of the supernatant, the cells were rinsed twice with ice-cold PBS and harvested using a plastic scraper.

Preparation of a standard protein mixture and biotinylated membrane fractions

The standard proteins, human apo-transferrin, human fetuin, bovine ribonuclease B (RNase B), chicken ovalbumin, and bovine serum albumin (BSA) were dissolved with a solution of 10 mM sodium periodate in cold PBS (pH 7.2), and then incubated in the dark at 4 °C for 30 min. The total amount of five proteins were 20 μg. Glycerol was added to a final concentration of 15 mM, and the mixture was incubated for 5 min on ice or at 4 °C to stop the oxidation reaction. The buffer was exchanged with PBS using a Bio-Spin chromatography column from Bio-Rad Labs (Richmond, CA, USA). The standard protein mixture was then incubated with 200 μg/mL EZ-Link Biotin-LC-Hydrazide solution in PBS (pH 7.2) for 1 h. The biotinylated proteins were digested with sequence grade trypsin. Half of these tryptic peptides were enriched by affinity capture, and the other half were directly treated with PNGase F as the control.

Biotinylated hepatocellular carcinoma cells were collected by centrifugation at 350 g for 10 min, resuspended in ice-cold hypotonic buffer (10 mM HEPES, pH 7.5, 1.5 mM $MgCl_2$, 10 mM KCl, 1× protease inhibitor cocktail, 1 mM NaF, and 1 mM Na_3VO_4), and incubated on ice for 15 min. The final amount of protease inhibitor cocktail in hypotonic buffer is 1 tablet per 10 ml of the solution. The cells were broken by Dounce homogenization (50 passes, homogenizer from Kontes Glass Co., Vineland, NJ). The cell lysates were then centrifuged at 1,000 g for 10 min to remove large cell debris and nuclei. The supernatant (the membrane fraction) was further centrifuged for 2 h at 100,000 g, and the pellet of the microsomal fraction was collected for subsequent analysis. All centrifugations were carried out at 4 °C.

Protein extraction and tryptic digestion of the membrane fractions

A two stage sequential extraction method was adopted for membrane fractions. First, the pellet was solubilized with 8 M urea in 400 mM NH_4HCO_3, pH 8.5. After centrifugation at 20,000 g for 10 min at 4 °C, the supernatant was collected. Then, 60 % methanol/20 mM NH_4HCO_3 solution was added to dissolve the pellet that was not solubilized by 8 M urea buffer, and the supernatant was collected by centrifugation as described above. The protein concentration was determined by Bradford method. Protein samples extracted by 8 M urea buffer were reduced with 8 mM TCEP at room temperature for 30 min, alkylated with 12 mM iodoacetamide for 30 min in darkness, and diluted 8-fold with 20 mM NH_4HCO_3. Protein samples were then digested with sequence grade trypsin at an enzyme:substrate ratio of 1:100 (w/w) at 37 °C for 18 h. The digest solution was passed over a C_{18} column (Handee spin column from Pierce (Rockford, US) packed with the C_{18} reversed-phase resin (SP-ODS-AP, 5 μm, 120-Å pore size) from Jinouya, Beijing) to remove extra urea and NH_4HCO_3. Tryptic peptides were eluted from the column with 80 % acetonitrile in 0.1 % TFA and dried in a Speed-Vac concentrator (Thermo). The protein samples extracted by 60 % methanol buffer were treated according to the work of Blonder *et al* [18], in which the protein samples were reduced and alkylated, followed by tryptic digestion in the same buffer (60 % methanol/20 mM NH_4HCO_3). The tryptic peptides from the sequential extraction were combined.

Avidin affinity enrichment of biotinylated glycopeptides

Biotinylated glycopeptides were enriched from the tryptic digestion by avidin affinity chromatography as follows. Dried tryptic peptides were first dissolved in affinity loading buffer (2× PBS,pH 7.2) and applied to a column packed with immobilized monomeric avidin (ABI) pretreated with 30 % acetonitrile with 0.4 % trifluoroacetic acid(TFA) and equilibrated with 2× PBS,pH 7.2. After sequential washing with (a) 2× PBS, pH 7.2, (b) 1× PBS,pH 7.2, (c) 50 mM NH_4HCO_3 containing 20 % methanol, pH 8.3, and (d) distilled water, the bound peptides were eluted with 30 % acetonitrile in 0.4 % TFA and concentrated using a Speed-Vac concentrator.

PNGase F treatment

Glycopeptide-enriched fractions were resuspended in 40 μL 20 mM NH_4HCO_3 and treated with 0.2 units of PNGase F (Roche). The enzymatic reaction was incubated at 37 °C overnight. Deglycosylated fractions were evaporated to dryness and resuspended in an aqueous solution that contained 0.1 % TFA.

Mass spectrometry analysis

The peptides from the standard protein sample were analyzed by MALDI-TOF MS, which was performed on a 4,800 Proteomics Analyzer MALDI-TOF/TOF instrument (Applied Biosystems). After desalting with the C_{18} ZipTip from

Millipore (Bedford, MA, US), all samples were mixed at a ratio of 1:9 with 5 mg/mL α-cyano-4-hydroxycinnamic acid in 50 % acetonitrile supplemented with 0.1 % TFA, and a 0.5 μL sample was applied to the MALDI target plate. Prior to analysis, the mass spectrometer was externally calibrated with 7 peptides obtained from a tryptic digest of myoglobin. The m/z range of the MS scan was from 600 to 4,000. Mass spectra were acquired in the positive reflector mode.

The peptides from the membrane fraction of hepatocellular carcinoma cells were analyzed by nano-LC-ESI-MS/MS, which was carried out on an LC-ESI-MS/MS system with an Agilent 1,100 series LC coupled with a hybrid linear ion trap-7 T Fourier transform ion cyclotron resonance mass spectrometer (LTQ-FTMS). The sample was loaded by a loading pump onto a fused silica capillary column prepared in-house with C_{18} packing materials (SP-ODS-AP, 5 μm, 120 Å pore size, Jinouya, Beijing). The mobile phase consisted of: A, 0.1 % formic acid/2 % acetonitrile, and B, 0.1 % formic acid/80 % acetonitrile. A 90 min linear gradient from 0 % B to 45 % B was applied at a flow rate of 200 nL/min. For each cycle, one full MS scan [400–2,000 m/z, 10^5 resolution in ICR cell and automatic gain control (AGC) of 5×10^5] was followed by 10 data-dependent MS/MS spectra in the linear ion trap from the 10 most abundant ions, with dynamic exclusion for 30 s. Singly charged ions were rejected.

Database searching and data interpretation

MS/MS spectra were searched with the Mascot™ search algorithm (Version 2.1, Matrix Science, London, UK) against a composite database containing the IPI human database (v.3.28) and its reverse database. Merge files were generated from the raw data format using the open source software DTAsupercharge (http://msquant.sourceforge.net/). Parameters for the database search were as follows. Variable modifications were carbamidomethyl for cysteines, oxidation for methionines, and an asparagine to aspartic acid conversion of +0.984 Da. The maximum number of missed cleavage sites was two. The precursor mass tolerance was 10 ppm and the fragment mass tolerance was 0.6 Da. Spectra with a Mascot score >32 (significance threshold $p<0.05$) and a valid glycosylation consensus sequence N-X-S/T were considered for further manual evaluation. The Uniprot system (http://www.uniprot.org/uniprot) was used for glycosylation data evaluation and protein function analysis. Assignment of protein cellular localization was according to the Uniprot system and GO (gene ontology) annotation. The TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) and SOSUIsignal (http://bp.nuap.nagoya-u.ac.jp/sosui/sosuisignal/sosuisignal_submit.html) programs were used to predict transmembrane domains and signal sequences.

Results and discussion

The affinity capture approach

The workflow of affinity capture approach is outlined in Fig. 1. First, intact cells were incubated with sodium periodate to oxidize the vicinal hydroxyl groups on the glycan of glycoproteins to aldehyde groups. Excess periodate was washed away from the cells to avoid oxidizing other subcellular components after cell lysis. The biotin-LC-hydrazide reagents were then added to react with aldehyde groups of cell surface glycoproteins, labeling the membrane glycoproteins with biotin groups. After lysis in ice-cold hypotonic buffer by Dounce homogenization and centrifugation to remove large debris, the cell surface membrane components in the cell lysate were fractionated by a single ultracentrifugation step. The extracted proteins were then cleaved into peptides by tryptic digestion, and the biotinylated glycopeptides were captured by immobilized avidin. Finally, the samples were treated with PNGase F to remove glycans and analyzed by LC-ESI-MS/MS.

To demonstrate the feasibility of the biotin label and the glycoprotein capture efficiency, we first tested the method with a standard glycoprotein mixture containing four glycosylated proteins and one non-glycosylated protein: human apo-transferrin, human fetuin, bovine ribonuclease B (RNase B), chicken ovalbumin, and BSA. The protein mixtures were oxidized, labeled with biotin, digested, enriched by affinity chromatography with immobilized avidin, deglycosylated with PNGase F and analyzed by MALDI-TOF/TOF. In total, five of the seven N-glycosylation sites previously reported from the four glycoproteins were unambiguously identified, and most of the non-glycosylated peptides were effectively eliminated (Fig. 2). Three replicates of the enriched peptide mixture were analyzed by MS to avoid false negatives. Details are described in Supplementary Table 1 in the Supporting Information. One glycosylation site on RNase B and one on ovalbumin were not directly identified with this method. One possible reason that the glycopeptide from RNase B was not detected is that the molecular weight of the deglycosylated glycopeptide (lower than 400 Da) is too low to be detected by MS. Conversely, the glycopeptide of ovalbumin may not have been detected as the tryptic peptide was too large (above 3,000 Da) to be sufficiently ionized and was lysine-terminated with less sensitive detection by MALDI-TOF MS [12, 19]. We analyzed the enriched peptide mixture with nano-LC-MS/MS and identified the missing glycopeptides (Supplementary Figure 1 in the Supporting Information), which confirms that the N-glycopeptides missed in MALDI-TOF MS analysis were also successfully labeled and captured.

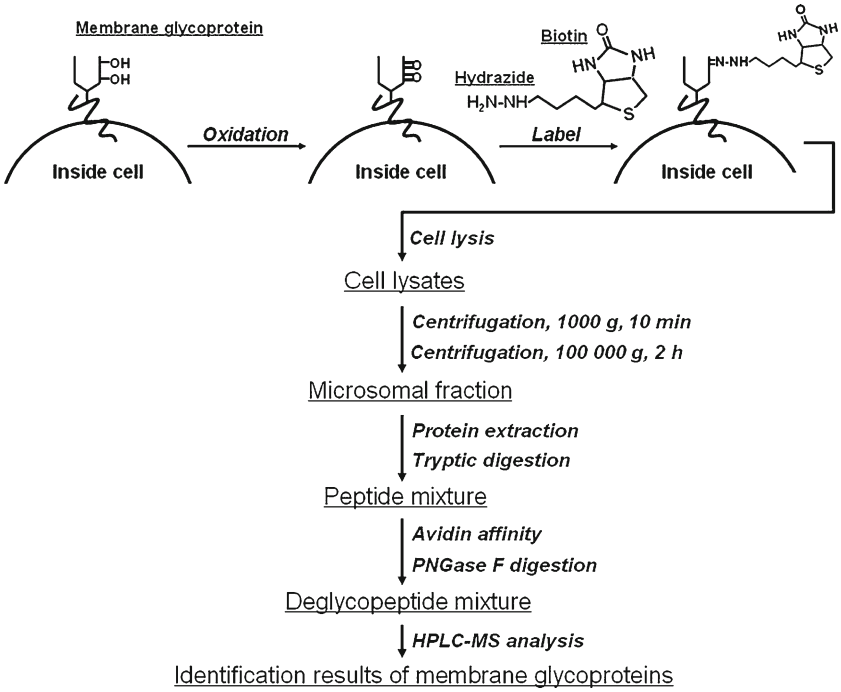

Fig. 1 Work flow diagram of the membrane glycoprotein affinity enrichment and identification procedure

Preparation of membrane glycoproteins

In this strategy, cell surface glycoproteins were labeled with the biotin group on the surface of intact cells, naturally separating them from other subcellular components. To avoid oxidizing and labeling these subcellular components after cell lysis, the reaction solutions were removed and the cells were washed with PBS several times after the oxidation and labeling reactions. By combining live cell surface labeling and avidin-affinity peptide enrichment, this method exclusively targets the set of cell surface-expressed glycoproteins. The affinity capture enrichment didn't depend on

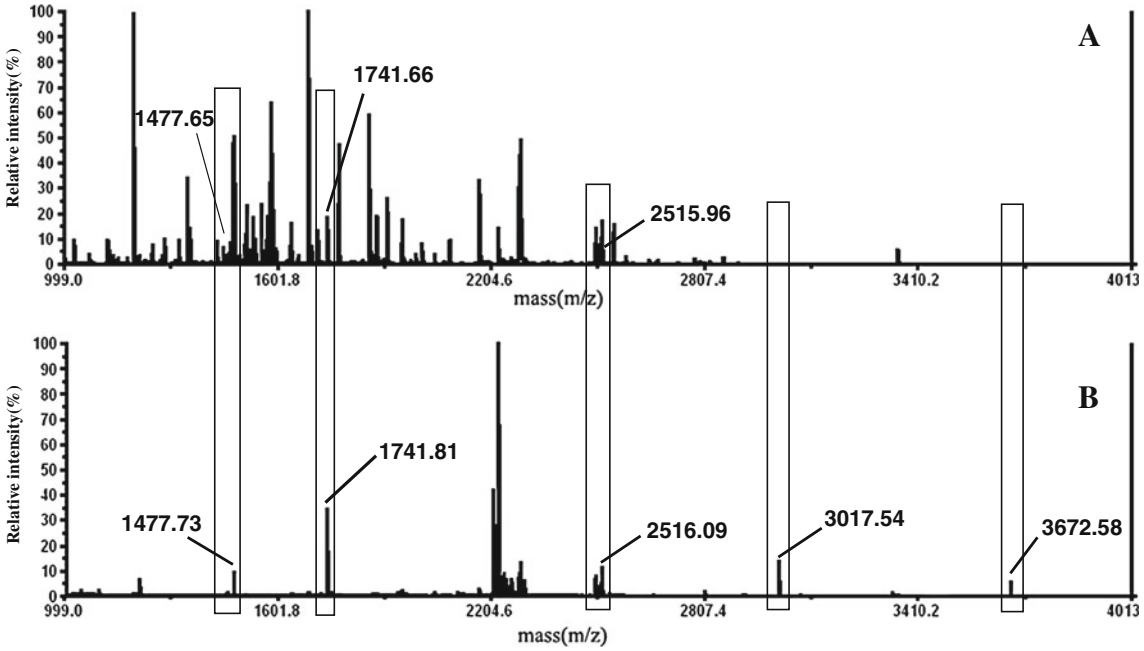

Fig. 2 The efficiency of the affinity method for enriching glycopeptides. MALDI-TOF mass spectra of peptide mixture from five standard proteins (**a**), and glycopeptides enriched from five standard proteins (**b**). The glycopeptides CGLVPVLAENYNK (1,477 m/z) and QQQHLFGSNVTDCSGNFCLFR (2,515 m/z) were from transferrin. The glycopeptides LCPDCPLLAPLNDSR (1,741 m/z), VVHAVEVALATFNAESNGSYLQLVEISR (3,017 m/z) and RPTGEVYDIEIDTLETTCHV LDPTPLANCSVR (3,672 m/z) were from fetuin. The strongest peak in **b** belonged to a peptide from trypsin produced by self-digestion (2,211 m/z,LGEHNIDVLEGNEQFINAAK)

Table 1 The list of identified glycoproteins and N- glycosylation sites on the surface of HCC cells

Uniprot	Protein Name	Glycopeptide	Sites	Score
P11279	lysosomal-associated membrane protein 1 (CD107)	K.N*MTFDLPSDATVVLNR.S	62	63
		R.GHTLTLN*FTR.N	103	34
		R.SSCGKEN*TSDPSLVIAFGR.G	84	83
		R.YSVQLMSFVYN*LSDTHLFPN*ASSK.E	121, 130	70
Q9UBG0	Macrophage mannose receptor 2 precursor (CD280)	K.KKPN*ATAEPTPPDR.W	364#	75
		R.TSN*ISKPGTLER.G	140#	44
		R.VTPACN*TSLPAQR.W	69#	62
P54709	Sodium/potassium-transporting ATPase subunit beta-3 (CD298)	K.FLKPYTLEEQKN*LTVCPDGALFEQK.G	124#	64
P05362	Intercellular adhesion molecule 1 (CD54)	R.AN*LTVVLLR.G	145	58
P13473	Isoform LAMP-2A of Lysosome-associated membrane glycoprotein 2 precur (CD107b)	K.IAVQFGPGFSWIAN*FTK.A	101	72
P13598	Intercellular adhesion molecule 2 (CD102)	K.AAPAPQEATATFN*STADR.E	176	45
P13987	CD59 glycoprotein precursor (CD59)	K.TAVN*CSSDFDACLITK.A	43	93
P27487	Dipeptidyl peptidase 4 (CD26)	K.LDFIILN*ETK.F	520	42
		R.IQN*YSVMDICDYDESSGR.W	321	34
Q5ZPR3	Isoform 2 of CD276 antigen precursor (CD276)	R.TALFPDLLAQGN*ASLR.L	104#	78
		R.VVLGAN*GTYSCLVR.N	215#	53
P35613	Isoform 2 of Basigin precursor (CD147)	K.ILLTCSLN*DSATEVTGHR.W	160	141
		K.ITDSEDKALMN*GSESR.F	268	95
P02786	Transferrin receptor protein 1 (CD71)	K.DFEDLYTPVN*GSIVIVR.A	251	87
		R.KQNNGAFN*ETLFR.N	727	47
Q10589	Bone marrow stromal antigen 2 precursor (CD317)	K.GFQDVEAQAATCN*HTVMALMASLDAEK.A	92#	91
		R.N*VTHLLQQELTEAQK.G	65#	60
P32004	Isoform 1 of Neural cell adhesion molecule L1 precursor (CD171)	R.LLFPTN*SSSHLVALQGQPLVLECIAEGFPTPTIK.W	247#	57
		R.THN*LTDLSPHLR.Y	979#	53
P06756	Isoform 1 of Integrin alpha-V precursor (CD51)	K.AN*TTQPGIVEGGQVLK.C	74#	61
		K.ISSLQTTEKN*DTVAGQGER.D	874#	57
		R.TAADTTGLQPILNQFTPAN*ISR.Q	615	79
P16070	Isoform 12 of CD44 antigen precursor (CD44)	K.AFN*STLPTMAQMEK.A	57#	64
P48960	Isoform 1 of CD97 antigen precursor (CD97)	K.TFKN*ESENTCQDVDECQQNPR.L	108#	51
P26006	Isoform Alpha-3A of Integrin alpha-3 precursor (CD49C)	K.N*ITIVTGAPR.H	265#	54
		R.SLDAYPILNQAQALEN*HTEVQFQK.E	605#	35
		K.LRPIIISMN*YSLPLR.M	573#	71
P08648	Integrin alpha-5 precursor (CD49)	K.GN*LTYGYVTILN*GSDIR.S	297#, 307#	66
		K.NLN*NSQSDVVSFR.L	773#	33
		R.VTGLN*CTTNHPINPK.G	868#	71
P32970	CD70 antigen (CD70)	R.GDTLCTN*LTGTLLPSR.N	170	38
Q6YHK3	Isoform 1 of CD109 antigen precursor (CD109)	K.TASN*LTVSVLEAEGVFEK.G	68	46
		R.GISDN*YTLALITYALSSVGSPK.A	1086#	109
P15144	Aminopeptidase N (CD13)	K.AEFN*ITLIHPK.D	234	82
P04156	Major prion protein (CD230)	K.GEN*FTETDVK.M	197	55
P08174	Isoform 1 of Complement decay-accelerating factor precursor(CD55)	K.GSQWSDIEEFCN*R.S	95#	73
P08195	4 F2 cell-surface antigen heavy chain (CD98)	K.DASSFLAEWQN*ITK.G	264	95
		K.SLVTQYLN*ATGNR.W	323	105
		R.LLIAGTN*SSDLQQILSLLESNK.D	280	89
P05556	Isoform Beta-1C of Integrin beta-1 precursor (CD29)	R.NPCTSEQN*CTSPFSYK.N	212#	81
		K.CHEGN*GTFECGACR.C	481#	42
		K.DTCTQECSYFN*ITK.V	669#	53

Table 1 (continued)

Uniprot	Protein Name	Glycopeptide	Sites	Score
		K.SCGECIQAGPNCGWCTN*STFLQEGMPTSAR.C	50#	66
		R.KEN*SSEICSNNGECVCGQCVCR.K	520#	76
Q9H5V8	Isoform 3 of CUB domain-containing protein 1 precursor (CD318)	R.IGTFCSN*GTVSR.I	180#	50
Q9Y4L1	150 kDa oxygen-regulated protein precursor	K.EN*GTDTVQEEEESPAEGSK.D	596	65
		R.AEPPLN*ASASDQGEK.V	931	71
		R.VFGSQN*LTTVK.L	515	62
P28300	LOX Protein-lysine 6-oxidase precursor	R.RRDPGAAVPGAAN*ASAQQPR.T	81	60
Q16790	Carbonic anhydrase 9	R.YFQYEGSLTTPPCAQGVIWTVFN*QTVMLSAK.Q	346	75
Q9UI10	Isoform 1 of Translation initiation factor eIF-2B subunit delta	K.VLLGAHALLAN*GSVMSR.V	411&	63
Q9NR50	Isoform 1 of Translation initiation factor eIF-2B subunit gamma	K.YIVDFLMEN*GSITSIR.S	218&	88
Q01650	Large neutral amino acids transporter small subunit 1	K.GDVSNLDPN*FSFEGTK.L	230#	42
		K.VASVININPN*TTHSTGSCR.S	257	62
Q13641	Trophoblast glycoprotein precursor	K.VLHN*GTLAELQGLPHIR.V	275#	90
		R.RPPLAELAALN*LSGSR.L	124#	57
Q16563	Synaptophysin-like protein 1	K.GQTEIQVNCPPAVTEN*K.T	71#	56
O00299	Chloride intracellular channel protein 1	K.GVTFN*VTTVDTK.R	42&	49
P07339	Cathepsin D precursor	K.GSLSYLN*VTR.K	263	43
P07602	Isoform Sap-mu-0 of Proactivator polypeptide precursor	K.DVVTAAGDMLKDN*ATEEEILVYLEK.T	80	65
		K.TCDWLPKPN*MSASCK.E	101	34
		R.TN*STFVQALVEHVKEECDR.L	215	50
P00533	Isoform 1 of Epidermal growth factor receptor precursor	K.DSLSIN*ATNIK.H	352	65
		K.EITGFLLIQAWPEN*R.T	413	37
		K.N*CTSISGDLHILPVAFR.G	361	74
		K.TCPAGVMGEN*NTLVWK.Y	603	66
		R.EFVENSECIQCHPECLPQAMN*ITCTGR.G	568	40
Q15758	Neutral amino acid transporter B	R.SYSTTYEERN*ITGTR.V	212#	39
P53396	ATP-citrate synthase	R.GKLGLVGVN*LTLDGVK.S	75&	67
P10606	Cytochrome c oxidase subunit 5B, mitochondrial precursor	R.IVGCICEEDN*TSVVWFWLHK.G	97&	49
P04062	Isoform Long of Glucosylceramidase precursor	R.DLGPTLAN*STHHNVR.L	309	58
P04114	Apolipoprotein B-100	R.FN*SSYLQGTNQITGR.Y	1523	35
		R.VNQNLVYESGSLN*FSK.L	2982	47
Q9BVK6	transmembrane emp24 protein transport domain containing 9	R.FTFTSHTPGEHQICLHSN*STK.F	104	38
Q08380	Galectin-3-binding protein precursor	R.ALGFEN*ATQALGR.A	69	37
		K.GLN*LTEDTYKPR.I	398	52
P98073	Enteropeptidase	R.N*CSIAGWGTVVYQGTTANILQEADVPLLSNER.C	909	37
P20645	Cation-dependent mannose-6-phosphate receptor precursor	R.EAGN*HTSGAGLVQINK.S	83	81
Q6FI81	Isoform 3 of Anamorsin	R.ILRPGGCLFLKEPVETAVDN*NSK.V	105&	43
P14625	Endoplasmin precursor	K.GVVDSDDLPLN*VSR.E	445	32
Q14126	desmoglein 2 preproprotein	K.DTGELN*VTSILDREETPFFLLTGYALDAR.G	112	55
		K.IN*ATDADEPNTLNSK.I	182	97
O60568	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 precursor	R.SAEFFN*YTVR.T	63#	52
Q04941	Proteolipid protein 2	RLSAPGCWAACTN*FSR.T	18	40
Q9HD45	Transmembrane 9 superfamily protein member 3 precursor	R.IVDVN*LTSEGK.V	174#	52

Table 1 (continued)

Uniprot	Protein Name	Glycopeptide	Sites	Score
Q9HDC9	Adipocyte plasma membrane-associated protein	R.AGPN*GTLFVADAYK.G	160	52
Q9Y277	Isoform 1 of Voltage-dependent anion-selective channel protein 3	K.VNN*ASLIGLGYTQTLRPGVK.L	239&	51
P26447	Protein S100-A4	K.LN*KSELKELLTR.E	30&	34
O95721	Synaptosomal-associated protein 29	K.SKPVETPPEQN*GTLTSQPNNR.L	135&	34
Q8N335	Glycerol-3-phosphate dehydrogenase 1-like protein	K.LPENVVAMSN*LSEAVQDADLLVFVIPHQFIHR.I	80&	70
Q32P28	Isoform 3 of Prolyl 3-hydroxylase 1 precursor	K.VPLQSAHLYYN*VTEK.V	540	40
Q96JJ7	Isoform 1 of Protein disulfide-isomerase TXNDC10 precursor	K.LVALAVIDEKN*TSVEHTR.L	258#	52
P10586	Receptor-type tyrosine-protein phosphatase F precursor	R.DINSQQELQN*ITTDTR.F	956	93
Q86Z14	Beta-klotho	K.NPNFTPVN*ESQLFLYDTFPK.N	71#	61
P21796	Voltage-dependent anion-selective channel protein 1	K.VNN*SSLIGLGYTQTLKPGIK.L	238&	63
P16403	Histone H1.2	K.ALAAAGYDVEKN*NSR.I	76&	51
Q13061	Triadin	MTEITAEGN*ASTTTTVIDSK.N	9&	38
P46977	Dolichyl-diphosphooligosaccharide—protein glycosyltransferase subuni	R.TILVDN*NTWN*NTHISR.V	544#, 548#	78
Q13433	Isoform 1 of Zinc transporter SLC39A6 precursor	R.NTNENPQECFN*ASK.L	260#	68
P22888	Lutropin-choriogonadotropic hormone receptor	R.IEANAFDNLLN*LSEILIQNTK.N	99#	37
P13637	Sodium/potassium-transporting ATPase alpha-3 chain	K.VDN*SSLTGESEPQTR.S	205&	35
O00469	Isoform 2 of Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 precur	R.YN*CSIESPR.K	696#	62
P62937	Peptidyl-prolyl cis-trans isomerase A	K.HTGPGILSMANAGPNTN*GSQFFICTAK.T	108&	75
Q96AY3	FK506-binding protein 10 precursor	R.TLSRPSETCN*ETTK.L	393#	54
Q86V97	Kelch repeat and BTB domain-containing protein 6	R.LDLTN*CTAILK.F	172&	38
P43308	Translocon-associated protein subunit beta precursor	R.IAPASN*VSHTVVLRPLK.A	88	96
P30475	HLA class I histocompatibility antigen, B-39 alpha chain	R.GYYN*QSEAGSHTIQR.M	110#	96
P23284	peptidylprolyl isomerase B precursor	K.DTN*GSQFFITTVK.T	148#	59
P05026	Isoform 2 of Sodium/potassium-transporting ATPase subunit beta-1	K.YLQPLLAVQFTN*LTMDTEIR.I	265#	90
Q15904	V-type proton ATPase subunit S1	K.QPVSPVIHPPVSYN*DTAPR.I	261#	39
O15031	similar to Plexin-B2 precursor	K.SCVAVTSAQPQN*MSR.R	528#	63

Glycosylation sites are marked with * according to the Uniprot database. & The glycosylation sites are not annotated in the database; #Potential glycosylation sites are annotated in the database

the purification of the plasma membrane, so only a low-speed centrifugation to remove cell debris and an ultra-high-speed centrifugation to collect the pellet of the microsomal fraction were needed. This is obviously superior to the most commonly used sub-cellular fractionation methods for isolating cell surface glycoproteins [13, 14], the efficacy of which greatly depends on the purity of the membrane and is always confounded by the non-cell surface proteins. Moreover, because of its simplicity, the approach is rapid. In our laboratory, we found that it is able to be completed a day faster than strategies based on membrane enrichment such as sucrose gradient centrifugation.

For the membrane protein extraction, a non-detergent urea solution and organic solvents (60 % methanol / 20 mM NH_4HCO_3) were used in sequence. Although detergents help to dissolve the membrane proteins, we did not use them in the extraction solution since detergents interfere with LC separation and severely suppress the ionization efficiency in MS analysis [20]. While 8 M urea solution increases the solubility of membrane proteins [21] and 60 %

methanol can destroy the membrane lipid bilayer [22], neither of them suppress the efficiency of mass spectrometry ionization or liquid separation. Compared with lysis in either 8 M urea or 60 % methanol alone, the two-stage extraction results in the near complete solubilization of membranes with no precipitated fraction observed in the lysis buffer [23].

Selective labeling of the membranes of hepatocellular carcinoma cells and identification of the captured glycopeptides

The strategy was applied to the surface glycoproteomic analysis by using live hepatocellular carcinoma cells. In total, 116 non-redundant glycosylated peptides were identified, corresponding to 119 glycosylation sites and 79 non-redundant glycoproteins. Of the 79 identified proteins, more than two glycopeptides were identified in 22 of them, and one glycopeptide was identified in 57 of them.

Upon comparing the identified glycosylation sites with the Uniprot database (version of January 2009), 54 sites (45 %) have been reported previously, and 49 sites (41 %) were described as potential glycosylation sites. To our knowledge, this study provides the first experimental evidence for these 49 N-glycosylation sites. 16 sites were novel sites not in the database. Thus, more than half of the N-glycosylation sites identified in our study were either listed as potential sites or were not identified in the Uniprot database. Table 1 provides the details of the identified N-linked glycosylation sites and the identity of their parent glycoproteins. The reliable identification of peptides and N-glycosylation sites were based on the following principles: (a) All the identified peptides must have a Mascot ion score above 32 ($P<0.05$) and their expected value must be evaluated. (b) All identified glycoproteins must have at least one peptide containing the consensus N-X (not P)-S/T glycosylation motif. (c) All identified glycopeptides must have an asparagine to aspartic acid deamidation site with a mass increase of 0.984 Da. (d) All the MS/MS spectra for each identified glycopeptide must be manually checked. Figure 3 shows an example of the MS/MS spectrum of the N-linked glycopeptide TAVN^CSSDFDACLITK from the precursor of protein CD59. The mass difference of 115 Da between fragments y12 and y13 confirms that a conversion of N (asparagine) to D (aspartic acid) occurred at site Asn-43.

The capture efficiency, specificity and reproducibility of the method for the hepatocellular carcinoma cells were evaluated. First, we checked the reproducibility of the method. With a significance threshold of $p<0.05$, a total of 142 non-redundant peptides were identified with 102,107 and 99 peptides in subsequent experiments. The reproducibility was 72.3 % on average, with separate identification ratios of 71.8 % (102/142), 75.3 % (107/142) and 69.7 % (99/142).

In addition, we also assessed the selectivity of the method for glycopeptides based on the number of glycopeptides compared to the total number of identified peptides. Based on the number of glycopeptides identified, the specificity of this method is 81.6 % (116/142). This proved that the method was effective, specifically targeted membrane glycoproteins, and was capable of removing non-glycopeptides from the sample mixture, thus reducing the sample complexity.

We classified the subcellular location of identified proteins according to Uniprot and GO annotation. Among the identified glycoproteins, 82 % (65) were classified as membrane proteins (Fig. 4a). Of these membrane proteins, 77 % (50) were localized to the plasma membrane and 12 % (8) were localized to the endoplasmic reticulum (Fig. 4b). Further details of the cellular localization are presented in Supplementary Table 2 in the Supporting Information. The identified cell surface membrane proteins included cell adhesion proteins, cell surface receptors and many solute carrier transporters. Some of these proteins were abundant housekeeping proteins, such as histones, which could potentially be cell surface components [24]. Other proteins (*e.g.*, Isoform 2 of Procollagen-lysine) could become components of the extracellular matrix once secreted [25], while yet others (*e.g.*, Cathepsin D) may have remained associated with the cells or extracellular matrix through protein-protein interactions [26, 27].

The identified glycoproteins of hepatocellular carcinoma cells were analyzed on the basis of the calculated average GRAVY (grand average of hydrophobicity) values and transmembrane domains (TMDs) were predicted to assess the efficacy of the developed protocol for the identification of membrane proteins (Supplementary Table 2 in the Supporting Information). Because the affinity approach in this study uses non-gel based separation, it is especially suitable for the purification of the highly hydrophobic integral membrane proteins. Numerous proteins with multiple membrane-spanning domains were detected: 68 % (54) of the identified glycoproteins had transmembrane domains (TMDs) and 24 % (19) of the proteins were predicted to contain 2–13 TMDs. The SOSUIsignal program is a high performance system for signal peptide prediction [28]. The glycoproteins with no detectable TMDs were further analyzed by the SOSUIsignal program, and 16 % (13) proteins were predicted to have a signal peptide, suggesting that these might be secreted proteins located on the surface cell membrane [29]. The predicted signal sequences of the proteins are reported in Supplementary Table 2 in the Supporting Information. In all, the potential integral membrane proteins with TMDs and secreted proteins with signal peptides accounted for 84 % (67) of total identified glycoproteins. In addition, 6 identified glycoproteins with no TMDs, including CD55, were predicted as membrane proteins by database annotation,

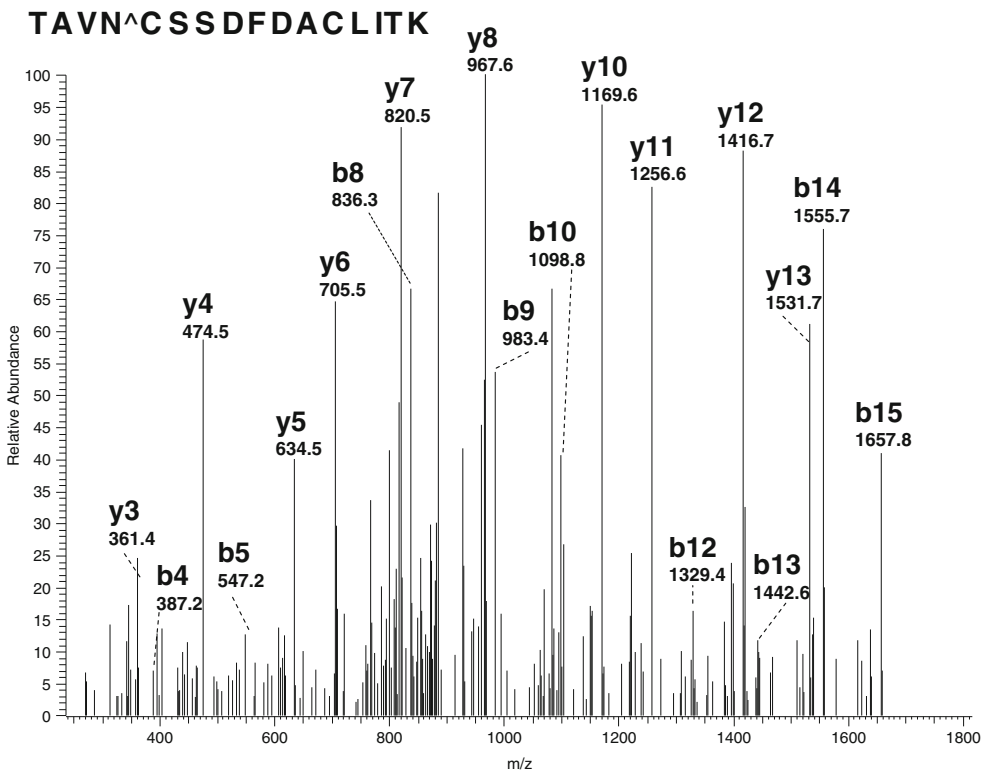

Fig. 3 MS/MS spectrum of the N-linked tryptic peptide TAVN^CSSDFDACLITK from CD59 glycoprotein precursor. Asn-43 has been converted to Asp by PNGase F as shown by the mass difference of 115 Da between fragments y_{12} and y_{13}

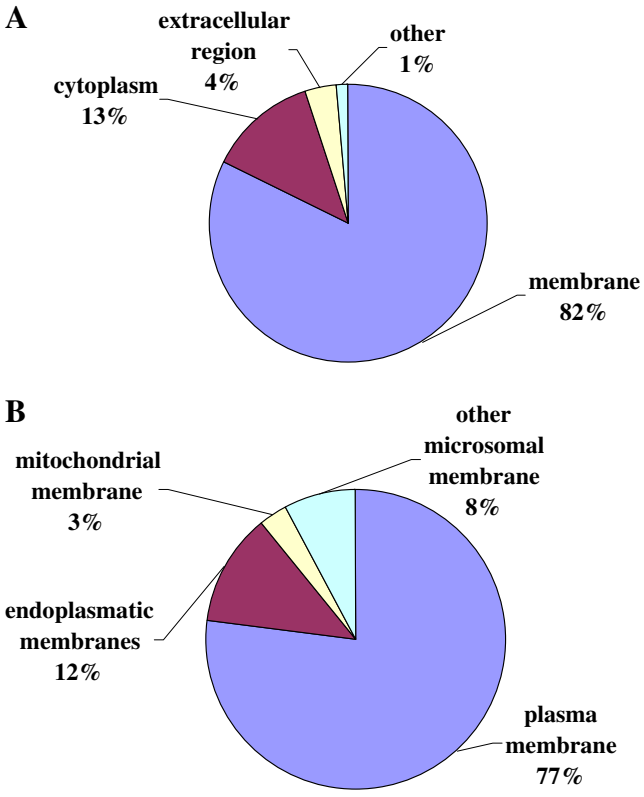

Fig. 4 Subcellular location of the glycoproteins determined by affinity enrichment and identification. (**a**) Total identified glycoproteins; (**b**) Membrane glycoproteins. All identified proteins were categorized using the UniProt database and GO annotation

and one of these proteins was predicted to be extracellular. These results further support the effectiveness of the method, which successfully identified integral membrane proteins containing multiple TMDs and showed no bias against the complex multispanning proteins. The GRAVY index of identified glycoproteins shows a broad scattering of values ranging from hydrophilic proteins such as Triadin (−1.31) to extremely hydrophobic proteins such as Proteolipid protein 2 (0.765). However, these two proteins both localized on plasma membrane and have TMDs despite the significant deviation in their GRAVY values. In general, there appeared to be a correlation between TMDs and the GRAVY values of the identified glycoproteins (Fig. 5). In the region with multiple TMDs (more than 3), GRAVY values tended to be positive and the proteins were predicted to be hydrophobic. However, in the region with 0–2 TMDs, the GRAVY values were distributed in the range of −0.4~0.4. According to the study of Santoni *et al* [30], it is difficult to determine whether a protein is hydrophobic or hydrophilic based on the GRAVY value of a protein distributed in this range.

The above results support the efficacy of the method in labeling the cell surface of hepatocellular carcinoma cells and identifying the captured glycopeptides with high efficiency and specificity. As only one protease was used in our peptide digestion, our rate of identification of glycoproteins could be improved by using a multiple enzyme strategy. Through the statistical analysis of the mass distribution of tryptic deglycopeptides of membrane glycoproteins in the Uniprot database (Human IPI3.28), the molecular weights

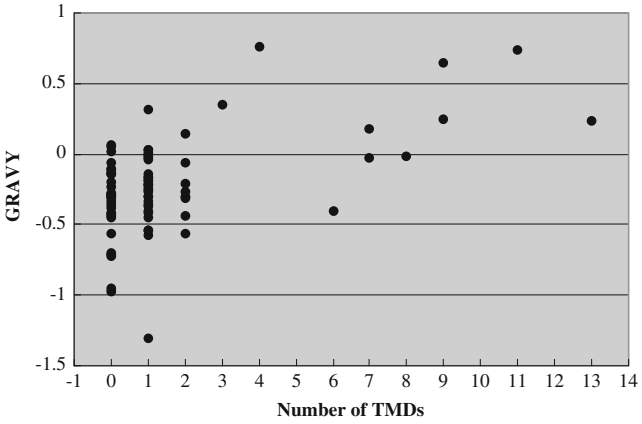

Fig. 5 The number of transmembrane domains (TMDs) of identified glycoproteins plotted against their GRAVY values

of 68 % of the deglycopeptides are between 800 Da and 3,500 Da, the ideal range for MS analysis. As tryptic digestion may result in a large number of glycopeptides with sizes outside this ideal range and may eventually affect the yields of identification, the use of several kinds of proteases may aid in larger scale identification [31].

Functional characterization of the identified membrane glycoproteins of hepatocellular carcinoma cells

The molecular functions of the membrane glycoproteins of hepatocellular carcinoma cells identified in this study were classified according to the GO database and literature surveys. A large number of these proteins are involved in inflammation, cell adhesion, signal transduction and other physiological processes. The identified glycoproteins were primarily analyzed by their known importance in cancer onset and progression.

One indicator of the effectiveness of our strategy for isolating surface antigens is the number of CD antigens identified in our analysis. A total of 26 CD antigens with 50 glycopeptides were detected in the membrane glycoproteins of hepatocellular carcinoma cells, comprising 43 % of the total glycopeptides identified (Table 1), which would be one of the potential uses in future studies. Many of these CD antigens identified have important biological functions and are prospective drug targets and biomarkers. CD44 is known to be involved in cell migration, tumor growth and progression [32]. It has many isoforms that are expressed in a cell-specific manner and are differentially glycosylated [33]. CD44 functions as an extracellular matrix receptor involved in cell-cell and cell-matrix interactions and may also participate in growth regulation by presenting growth factors to their cell surface receptors [34]. The interaction of CD44 with the cytoskeleton and various signaling molecules plays a pivotal role in promoting invasive and metastatic specific tumor phenotypes such as MMP-mediated matrix degradation, tumor cell growth, migration, and invasion. The abnormal expression of CD44 molecules can significantly increase the ability of HCC to metastasize [35–38]. The isoforms of CD44 are involved not only in tumor invasion and metastasis, but also in tumor immune escape [39]. The study of the modification of CD44 by glycosylation is important for understanding the role of CD44 and the mechanism of cancer metastasis. The only reported glycosylation site of CD44 in the Uniprot database was successfully identified by this affinity capture method.

CD147 is highly expressed in HCC and promotes tumor invasion and metastasis through the induction of matrix metalloproteinase (MMP) [40, 41]. The CD147 molecule has an extracellular domain containing 3 consensus sequences for potential glycosylation sites. There are 2 glycosylation sites and 1 potential glycosylation site in the Uniprot database. Here 2 validated glycosylation sites were identified. The degree of glycosylation of CD147 was reported to be tissue specific, as differences in glycosylation corresponding to different functions are found in different organs [42]. The glycosylation of the CD147 molecule determines its ability to activate MMP, and non-glycosylated CD147 cannot activate the secretion of MMP [43]. CD147 has been reported to serve as a potential pathological target for early detection of HCC [44].

A series of integrin proteins was identified including integrin alpha-5 precursor (CD49), isoform alpha-3A of integrin alpha-3 precursor (CD49C), isoform 1 of integrin alpha-V precursor (CD51) and isoform beta-1C of integrin beta-1 precursor (CD29). The expression level of integrin molecules in tumors often varies significantly [45], and these changes are conducive to the infiltration and metastasis of tumor cells [46]. A study of integrin glycosylation is not only beneficial to the understanding of the mechanism of liver cancer, but is also helpful for distinguishing the stage of hepatocellular carcinoma carcinogenesis. The expression level of integrins and their distribution changes with the progression of HCC cells through different stages of carcinoma. Different integrin molecules in different tissues and cells have particular patterns of distribution, which can be used as references for the diagnosis of HCC and can assist the classification of the clinical stage of HCC [47].

We also detected other classes of surface membrane proteins with critical biological functions associated with development and cancer. Epidermal growth factor receptor (EGFR) is a transmembrane glycoprotein, and its intracellular region has tyrosine kinase activity [48]. There are 11 glycosylation sites reported in Uniprot database for this protein, and we identified 5 sites including 2 partially glycosylated modification sites that may be related to the function of the protein [49, 50]. The interaction between EGF and EGFR ultimately leads to cell proliferation. There is

evidence that the EGFR tyrosine kinase receptor family plays an important role in cancer occurrence and development. High expression of EGFR promotes tumor cell proliferation, adhesion, invasion and metastasis, and inhibits tumor cell apoptosis. In liver cancer, the high expression of EGFR may be one of the causes of hepatic proliferation. Today, therapy targeting EGFR has become an effective method of tumor treatment [51, 52].

MHC-I molecules are transmembrane glycoproteins that initiate the immune response [53]. In addition to the pivotal role of MHC-I molecules in immune recognition, the MHC-I molecules also play a key role in preventing tumor progression and provide an important biomarker for monitoring the effect of drug treatment [54]. Our analysis identified an MHC-I glycosylation site that had not been reported in the Uniprot database.

Most transport proteins identified in this study were Na^+/K^+-ATP enzymes, including the sodium/potassium-transporting ATPase alpha-3 chain and isoform 2 of sodium/potassium-transporting ATPase subunit beta-1. Na^+/K^+-ATP enzymes transport Na^+/K^+ ions through the catalytic hydrolysis of ATP [55]. Na-K-ATPase can not only transport potassium ions, but also transduce signals. Na^+/K^+-ATP enzymes may activate a variety of cell signaling pathways, such as cell proliferation, apoptosis and other physiological and pathological processes [56, 57].

Future studies using affinity approach on hepatocellular carcinoma cells will seek to optimize and integrate it with various experimental workflows. For example, the strategy could be used to discover new and promising hepatocellular carcinoma biomarkers in combination with quantitative proteomic approaches, such as stable isotope labeling. By quantitatively analyzing the differential expression of cell surface glycoproteins, for example during different stages of hepatocellular carcinoma, and by identifying and authenticating them, this method could provide a catalog of new potential biomarkers correlated with cancer development as well as future drug targets. Another application of this method is to focus on the differentiation markers during cellular activation and differentiation [58]. A new capture microarray could also be developed that would have clinical applications in the future. The microarrays would contain streptavidin immobilized on the 96 micro-wells, and the microarray format would allow for rapid analysis of the binding of targeted biotin labeling glycoproteins of the cells coupled with mass spectrometry in a high-throughput fashion, potentially realizing process automation.

Conclusions

This work reports a glycoproteomic analysis of N-glycosylation sites on the cell surface proteins of human hepatocellular carcinoma (HCC) cells using a purification approach that involves biotinylation of cell surface glycoproteins, affinity enrichment of these glycoproteins and their identification by mass spectrometry. In our results, 82 % of the glycoproteins identified were classified as membrane proteins, 68 % had transmembrane domains (TMDs), and 24 % proteins were predicted to contain 2–13 TMDs. This approach increased the relative amount of membrane glycopeptides by trapping the glycopeptides and removing the non-glycopeptides from the sample, which reduced the complexity of the peptide mixture, eliminated the suppression of non-glycopeptides and benefited the subsequent identification of glycopeptides using mass spectrometry. Using stringent criteria, we identified 119 different N-glycosylation sites within 79 different glycoproteins. Among these, 65 N-glycosylation sites not reported in the Uniprot database have been identified based on our experimental evidence. This study also employed the developed affinity approach for the functional analysis of membrane glycoproteins of hepatocellular carcinoma (HCC) cells. Many proteins have been reported as potential biomarkers of hepatocellular carcinoma. These findings demonstrate the utility of the enrichment strategy for the global screening of membrane glycoproteins from complex biological samples. The identified proteins thus comprise the first database of glycoproteins from the membranes of HCC cells. The research provides an approach for the surface glycoproteomic analysis of HCC cellsand facilitates further researches on the discovery of potential biomarkers and therapeutic targets.

Acknowledgments We thank Dr. Weijie Qin and Dr. Zhishang Hu for their helpful advice and critical reading of the manuscript. This study was supported by the National Key Program for Basic Research of China (No. 2011CB910603), the International Scientific Cooperation Project of China (No. 2011DFB30370), the National Natural Science Foundation of China (No. 30900258, 20905077,31100591).

References

1. Wallin, E., von Heijne, G.: Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci. **7**(4), 1029–1038 (1998)
2. Jang, J.H., Hanash, S.: Profiling of the cell surface proteome. Proteomics **3**(10), 1947–1954 (2003)
3. Fauman, E.B., Hopkins, A.L., Groom, C.R.: Structural bioinformatics in drug discovery. Meth. Biochem. Anal. **44**, 477–497 (2003)
4. Holland, N.B., Qiu, Y., Ruegsegger, M., Marchant, R.E.: Biomimetic engineering of non-adhesive glycocalyx-like surfaces using oligosaccharide surfactant polymers. Nature **392**(6678), 799–801 (1998)
5. Gahmberg, C.G., Tolvanen, M.: Why mammalian cell surface proteins are glycoproteins. Trends Biochem. Sci. **21**(8), 308–311 (1996)

6. Hirabayashi, J., Kasai, K.: Separation technologies for glycomics. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. **771**(1–2), 67–87 (2002)
7. Hakomori, S.: Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. Cancer Res. **56**(23), 5309–5318 (1996)
8. Haltiwanger, R.S.: Regulation of signal transduction pathways in development by glycosylation. Curr. Opin. Struct. Biol. **12**(5), 593–598 (2002)
9. Parkin, D.M., Bray, F., Ferlay, J., Pisani, P.: Global cancer statistics, 2002. CA Cancer J. Clin. **55**(2), 74–108 (2005)
10. Sakamoto, M., Mori, T., Masugi, Y., Effendi, K., Rie, I., Du, W.: Candidate molecular markers for histological diagnosis of early hepatocellular carcinoma. Intervirology **51**(Suppl 1), 42–45 (2008)
11. Wu, C.C., Yates 3rd, J.R.: The application of mass spectrometry to membrane proteomics. Nat. Biotechnol. **21**(3), 262–267 (2003)
12. Sun, B., Ranish, J.A., Utleg, A.G., White, J.T., Yan, X., Lin, B., Hood, L.: Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. Mol. Cell Proteomics **6**(1), 141–149 (2007)
13. Ghosh, D., Krokhin, O., Antonovici, M., Ens, W., Standing, K.G., Beavis, R.C., Wilkins, J.A.: Lectin affinity as an approach to the proteomic analysis of membrane glycoproteins. J. Proteome Res. **3** (4), 841–850 (2004)
14. Zhang, H., Li, X.J., Martin, D.B., Aebersold, R.: Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat. Biotechnol. **21**(6), 660–666 (2003)
15. Pasquali, C., Fialka, I., Huber, L.A.: Subcellular fractionation, electromigration analysis and mapping of organelles. J. Chromatogr. B: Biomed. Sci. Appl. **722**(1–2), 89–102 (1999)
16. Roffman, E., Wilchek, M.: Selective covalent modification of newly synthesized activation-induced sialoglycoconjugates on the T lymphocyte membrane. Biochem. Biophys. Res. Commun. **135** (2), 473–479 (1986)
17. Roffman, E., Meromsky, L., Ben-Hur, H., Bayer, E.A., Wilchek, M.: Selective labeling of functional groups on membrane proteins or glycoproteins using reactive biotin derivatives and 125I-streptavidin. Biochem. Biophys. Res. Commun. **136**(1), 80–85 (1986)
18. Blonder, J., Chan, K.C., Issaq, H.J., Veenstra, T.D.: Identification of membrane proteins from mammalian cell/tissue using methanol-facilitated solubilization and tryptic digestion coupled with 2D-LC-MS/MS. Nat. Protoc. **1**(6), 2784–2790 (2006)
19. Beardsley, R.L., Karty, J.A., Reilly, J.P.: Enhancing the intensities of lysine-terminated tryptic peptide ions in matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun. Mass Spectrom. **14**(23), 2147–2153 (2000)
20. Funk, J., Li, X., Franz, T.: Threshold values for detergents in protein and peptide samples for mass spectrometry. Rapid Commun. Mass Spectrom. **19**(20), 2986–2988 (2005)
21. Foster, L.J., Zeemann, P.A., Li, C., Mann, M., Jensen, O.N., Kassem, M.: Differential expression profiling of membrane proteins by quantitative proteomics in a human mesenchymal stem cell line undergoing osteoblast differentiation. Stem Cells **23**(9), 1367–1377 (2005)
22. Karsan, A., Blonder, J., Law, J., Yaquian, E., Lucas, D.A., Conrads, T.P., Veenstra, T.: Proteomic analysis of lipid microdomains from lipopolysaccharide-activated human endothelial cells. J. Proteome Res. **4**(2), 349–357 (2005)
23. Dormeyer, W., van Hoof, D., Mummery, C.L., Krijgsveld, J., Heck, A.J.: A practical guide for the identification of membrane and plasma membrane proteins in human embryonic stem cells and human embryonal carcinoma cells. Proteomics **8**(19), 4036–4053 (2008)
24. Watson, K., Gooderham, N.J., Davies, D.S., Edwards, R.J.: Nucleosomes bind to cell surface proteoglycans. J. Biol. Chem. **274**(31), 21707–21713 (1999)
25. Myllyla, R., Wang, C., Heikkinen, J., Juffer, A., Lampela, O., Risteli, M., Ruotsalainen, H., Salo, A., Sipila, L.: Expanding the lysyl hydroxylase toolbox: new insights into the localization and activities of lysyl hydroxylase 3 (LH3). J. Cell. Physiol. **212**(2), 323–329 (2007)
26. Diment, S., Leech, M.S., Stahl, P.D.: Cathepsin D is membrane-associated in macrophage endosomes. J. Biol. Chem. **263**(14), 6901–6907 (1988)
27. Ogier-Denis, E., Bauvy, C., Couvineau, A., De Stefanis, D., Isidoro, C., Codogno, P.: Localization and processing of glycosylphosphatidylinositol anchored cathepsin D. Biochem. Biophys. Res. Commun. **211**(3), 935–942 (1995)
28. Gomi, M., Sonoyama, M., Mitaku, S.: High performance system for signal peptide prediction: SOSUIsignal. Chem. Bio. Inform. J. **4**(4), 142–147 (2004)
29. Nunomura, K., Nagano, K., Itagaki, C., Taoka, M., Okamura, N., Yamauchi, Y., Sugano, S., Takahashi, N., Izumi, T., Isobe, T.: Cell surface labeling and mass spectrometry reveal diversity of cell surface markers and signaling molecules expressed in undifferentiated mouse embryonic stem cells. Mol. Cell Proteomics **4**(12), 1968–1976 (2005)
30. Santoni, V., Kieffer, S., Desclaux, D., Masson, F., Rabilloud, T.: Membrane proteomics: use of additive main effects with multiplicative interaction model to classify plasma membrane proteins according to their solubility and electrophoretic properties. Electrophoresis **21**(16), 3329–3344 (2000)
31. Chen, R., Jiang, X., Sun, D., Han, G., Wang, F., Ye, M., Wang, L., Zou, H.: Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. J. Proteome Res. **8**(2), 651–661 (2009)
32. Takahashi, K., Eto, H., Tanabe, K.K.: Involvement of CD44 in matrix metalloproteinase-2 regulation in human melanoma cells. Int. J. Cancer **80**(3), 387–395 (1999)
33. Ponta, H., Sherman, L., Herrlich, P.A.: CD44: from adhesion molecules to signalling regulators. Nat. Rev. Mol. Cell Biol. **4** (1), 33–45 (2003)
34. Kajita, M., Itoh, Y., Chiba, T., Mori, H., Okada, A., Kinoh, H., Seiki, M.: Membrane-type 1 matrix metalloproteinase cleaves CD44 and promotes cell migration. J. Cell Biol. **153**(5), 893–904 (2001)
35. Qin, L.X., Tang, Z.Y.: The prognostic molecular markers in hepatocellular carcinoma. World J. Gastroenterol. **8**(3), 385–392 (2002)
36. Qin, L.X., Tang, Z.Y.: Recent progress in predictive biomarkers for metastatic recurrence of human hepatocellular carcinoma: a review of the literature. J. Cancer Res. Clin. Oncol. **130**(9), 497–513 (2004)
37. Hirohashi, K., Yamamoto, T., Uenishi, T., Ogawa, M., Sakabe, K., Takemura, S., Shuto, T., Tanaka, H., Kubo, S., Kinoshita, H.: CD44 and VEGF expression in extrahepatic metastasis of human hepatocellular carcinoma. Hepatogastroenterology **51**(58), 1121–1123 (2004)
38. Gao, C., Guo, H., Downey, L., Marroquin, C., Wei, J., Kuo, P.C.: Osteopontin-dependent CD44v6 expression and cell adhesion in HepG2 cells. Carcinogenesis **24**(12), 1871–1878 (2003)
39. Runger, T.M., Klein, C.E., Becker, J.C., Brocker, E.B.: The role of genetic instability, adhesion, cell motility, and immune escape mechanisms in melanoma progression. Curr. Opin. Oncol. **6**(2), 188–196 (1994)
40. Zhang, Q., Zhou, J., Ku, X.M., Chen, X.G., Zhang, L., Xu, J., Chen, G.S., Li, Q., Qian, F., Tian, R., Wen, N., Chen, Z.N.: Expression of CD147 as a significantly unfavorable prognostic factor in hepatocellular carcinoma. Eur. J. Canc. Prev. **16**(3), 196–202 (2007)
41. Jia, L., Xu, H., Zhao, Y., Jiang, L., Yu, J., Zhang, J.: Expression of CD147 mediates tumor cells invasion and multidrug resistance in hepatocellular carcinoma. Canc. Investig **26**(10), 977–983 (2008)

42. Guo, H., Majmudar, G., Jensen, T.C., Biswas, C., Toole, B.P., Gordon, M.K.: Characterization of the gene for human EMMPRIN, a tumor cell surface inducer of matrix metalloproteinases. Gene **220**(1–2), 99–108 (1998)
43. Sun, J., Hemler, M.E.: Regulation of MMP-1 and MMP-2 production through CD147/extracellular matrix metalloproteinase inducer interactions. Cancer Res. **61**(5), 2276–2281 (2001)
44. Mamori, S., Nagatsuma, K., Matsuura, T., Ohkawa, K., Hano, H., Fukunaga, M., Matsushima, M., Masui, Y., Fushiya, N., Onoda, H., Searashi, Y., Takagi, I., Tagiri, H.: Useful detection of CD147 (EMMPRIN) for pathological diagnosis of early hepatocellular carcinoma in needle biopsy samples. World J. Gastroenterol. **13** (21), 2913–2917 (2007)
45. Liu, L.X., Jiang, H.C., Liu, Z.H., Zhou, J., Zhang, W.H., Zhu, A.L., Wang, X.Q., Wu, M.: Integrin gene expression profiles of human hepatocellular carcinoma. World J. Gastroenterol. **8**(4), 631–637 (2002)
46. Masumoto, A., Arao, S., Otsuki, M.: Role of beta1 integrins in adhesion and invasion of hepatocellular carcinoma cells. Hepatology **29**(1), 68–74 (1999)
47. Uotani, H., Yamashita, I., Nagata, T., Kishimoto, H., Kashii, Y., Tsukada, K.: Induction of E-selectin after partial hepatectomy promotes metastases to liver in mice. J. Surg. Res. **96**(2), 197–203 (2001)
48. Mottolese, M., Natali, P.G., Coli, A., Bigotti, G., Benevolo, M., Cione, A., Raus, L., Carapella, C.M.: Comparative analysis of proliferating cell nuclear antigen and epidermal growth factor receptor expression in glial tumours: correlation with histological grading. Anticancer. Res. **18**(3B), 1951–1956 (1998)
49. Sato, C., Kim, J.H., Abe, Y., Saito, K., Yokoyama, S., Kohda, D.: Characterization of the N-oligosaccharides attached to the atypical Asn-X-Cys sequence of recombinant human epidermal growth factor receptor. J. Biochem. **127**(1), 65–72 (2000)
50. Wu, S.L., Kim, J., Hancock, W.S., Karger, B.: Extended Range Proteomic Analysis (ERPA): a new and sensitive LC-MS platform for high sequence coverage of complex proteins with extensive post-translational modifications-comprehensive analysis of beta-casein and epidermal growth factor receptor (EGFR). J. Proteome Res. **4**(4), 1155–1170 (2005)
51. Llovet, J.M., Bruix, J.: Molecular targeted therapies in hepatocellular carcinoma. Hepatology **48**(4), 1312–1327 (2008)
52. Giannelli, G., Sgarra, C., Porcelli, L., Azzariti, A., Antonaci, S., Paradiso, A.: EGFR and VEGFR as potential target for biological therapies in HCC cells. Cancer Lett. **262**(2), 257–264 (2008)
53. Parham, P.: Virtual reality in the MHC. Immunol. Rev. **167**, 5–15 (1999)
54. Chen, W., Cai, M.Y., Wei, D.P., Wang, X.: Pivotal molecules of MHC I pathway in human primary hepatocellular carcinoma. World J. Gastroenterol. **11**(21), 3297–3299 (2005)
55. Takeuchi, A., Reyes, N., Artigas, P., Gadsby, D.C.: The ion pathway through the opened Na(+), K(+)-ATPase pump. Nature **456** (7220), 413–416 (2008)
56. Bortner, C.D., Gomez-Angelats, M., Cidlowski, J.A.: Plasma membrane depolarization without repolarization is an early molecular event in anti-Fas-induced apoptosis. J. Biol. Chem. **276**(6), 4304–4314 (2001)
57. Kassardjian, A., Kreydiyyeh, S.I.: JNK modulates the effect of caspases and NF-kappaB in the TNF-alpha-induced down-regulation of Na+/K+ATPase in HepG2 cells. J. Cell. Physiol. **216**(3), 615–620 (2008)
58. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., Aebersold, R., Watts, J.D.: Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. Nat. Biotechnol. **27**(4), 378–386 (2009)