was confirmed by negative staining using rabbit nonimmune serum as a primary antibody.

■ RESULTS

Enrichment of LMW Polypeptides by QUEST-MS Technology

For large-scale biomarker discovery focusing on LMW proteome, we established a rapid and effective enrichment method for <20 kDa proteins in plasma (Figure 1A), termed quick enrichment of small targets for mass spectrometry (QUEST-MS). The QUEST-MS procedure included denaturation, reduction, alkylation, and subsequent tandem reversed phase extraction of plasma samples, for which it took only 1.5 h per 96 samples. The dried QUEST-MS-purified samples are now ready for usual protein assays such as SDS-PAGE or ELISA, while they can also be subjected to further mass spectrometric analyses after trypsinization. Crude plasma and purified proteins at each stage were separated on 16% Tris-tricine gel and stained by silver to visualize the efficiency of LMW protein enrichment after the first to fourth reversed-phase extraction by the Oasis HLB plate(Figure 1B). The quantitative density chromatograms of five lanes on the stained gel illustrated effective exclusion of >20 kDa proteins at every solid-phase extraction step (Figure 1C). In particular, two purifications by the Oasis HLB plate resulted in >90% enrichment ratio of <20 kDa LMW proteins. The principle of such easy and clear-cut separation of small proteins was based on pouring away huge proteins larger than pore diameter (8 nm) and preventing elution of middle-range proteins (20−100 kDa) by excess absorption effect to hydrophobic residues on the Oasis polymer beads. Although the third or more repetitions of reversed-phase purification could provide higher purity of LMW proteins, we accepted twice-purified samples for further biomarker discovery and validation experiments from the perspective of retaining advantages in throughput and cost.

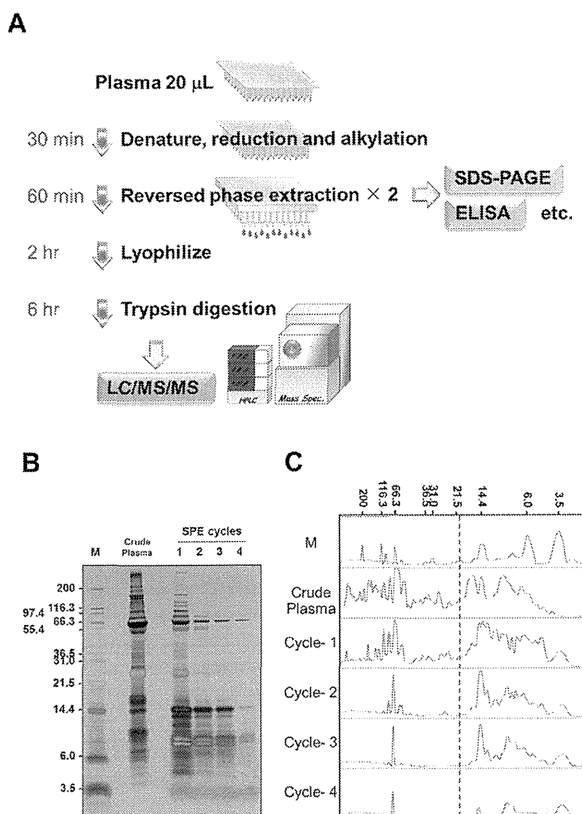


Figure 1. Quick enrichment of small targets for mass spectrometry (QUEST-MS) technology. (A) Schematic workflow of QUEST-MS purification is displayed. After twice so-phase extraction of denatured protein samples within 90 min, low-molecular-weight (LMW) proteins (<20 kDa) are enriched and ready for usual protein assays, such as SDS-PAGE or ELISA. The shotgun or targeted LC/MS/MS analyses can be performed following further tryptic digestion in typically 8 h. (B) Evaluation of LMW protein enrichment by silver-stained 16% Tris-tricine gel. M: molecular weight marker, SPE: solid phase extraction. (C) Result of quantitative densitometry analysis for the gel in panel B is shown. The vertical dashed line indicates 20.0 kDa. More than twice reversed phase purification achieved >90% enrichment of <20 kDa proteins.

LMW Biomarker Screening for Prostate Cancer

To avoid any sampling biases among institutes, we collected 116 plasma samples analyzed in the biomarker screening phase (Table 1) in a single hospital, employing strictly controlled standard operation procedures. The 20 μL of each plasma sample underwent one freeze−thaw step prior to QUEST-MS purification. From LC/MS/MS analysis of 116 purified samples, 153 057 nonredundant peptides were detected and quantified in the Expressionist RefinerMS module, as described in the Materials and Methods section. In particular, this LMW proteome catalog included four PSA-derived peptides $I_{25}VGGWECEK_{33}$ (m/z = 539.25, z = +2), $H_{34}SQPWQVLVASR_{45}$ (m/z = 469.92, z = +3), $A_{48}VCGGVLVHPQWVLTAAHCIR_{68}$ (m/z = 782.08, z = +3), and $L_{126}SEPAELTDAVK_{137}$ (m/z = 636.84, z = +2), which were detectable in 4, 3, 4, and 5 cases from the control group (healthy individuals + BPH patients), whereas they were detected in 15, 13, 10, and 11 cases from the prostate cancer group, respectively. This result suggested that general comparative tests based on the average of samples, such as *t* test or ANOVA, are not appropriate for the extraction of specific biomarkers from our data set because the valid values of low concentration tumor markers in the control group should be quite few. Therefore to maximize the specificity of new biomarker set, we employed the absent-present search algorithm, which extracted peptides exhibiting all-or-nothing feature between two groups (healthy control + BPH vs prostate cancer, Figure 2A). This search resulted in identification of 189 peaks demonstrating no or one-case detection among 43 controls, while at least 11-case detection among 73 prostate cancer patients (Figure 2B). The extracted peaks above could be considered as candidate biomarkers, which would provide lower false positive rate in addition to similar or higher sensitivity for the diagnosis of prostate cancer patients compared with PSA.

We identified 1126 nonredundant plasma proteins (Supplementary Table S1 in the Supporting Information) from 116 screening samples using Sequest database search analysis (FDR < 0.01). The molecular-weight distribution of 1126 proteins is displayed in Supplementary Figure S3 in the Supporting Information. For qualitative assessment of purified product, 201 out of 1126 identified proteins were <20 kDa. Because quantitative evaluation by SDS-PAGE (Figure 1C) showed >90% enrichment of <20 kDa proteins, most of the 318

dx.doi.org/10.1021/pr400547s | J. Proteome Res. 2013, 12, 4497−4506

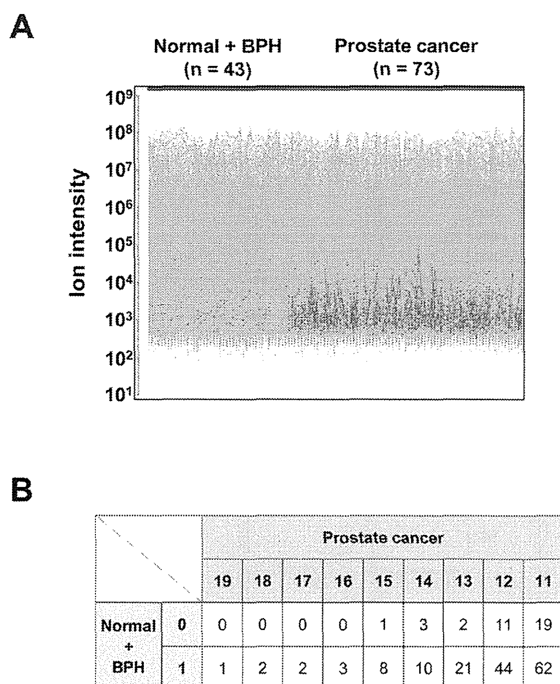


Figure 2. Prostate cancer biomarker screening by the absent/present search method on Expressionist proteome server. (A) 189 peptides out of 153 057 total peptides showed prostate cancer-specific detection patterns, demonstrating no or only one-case detection in the control group. The quantitative information of 189 peptides is indicated by red dots or lines on the 2D plane composed of 110 individuals as x axis and the peak intensity in LC/MS/MS as y axis. Normal: 24 healthy controls, BPH: 19 benign prostate hypertrophy patients. (B) Classification of 189 candidate biomarker peptides based on the frequency of detection. For instance, 62 peptides were detectable in 11 prostate cancer samples, while detected in 1 control sample (normal or BPH).

proteins more than 100 kDa (as UniProt database values) might exist in cleaved or degraded forms in purified product. Here, in addition to the detection of PSA, we observed low abundant chemokines (e.g., CCL14, CCL18, CXCL4, and CXCL7), hormones (e.g., gastric inhibitory polypeptide and inhibin β, whose concentrations were reported as 200−300 pg/mL[16] and 10−200 pg/mL,[17] respectively), hypotensive peptide adrenomedullin (50−100 pg/mL),[18] and others. By matching 189 candidate biomarker peptides against the imported protein identification data set on the Expressionist proteome server system, three peptides listed in Table 2 were successfully identified. From the aspect of highly specific mRNA expression in cerebral nervous system and prostate gland,[19] we subjected neuropeptide Y (NPY) to the next validation experiments.

NPY Expression in Prostate Cancer Tissues

NPY immunostaining patterns of prostate cancer frozen tissues at Gleason score 3 + 4 (Figure 3A,D) or 4 + 5 (Figure 3B,E)

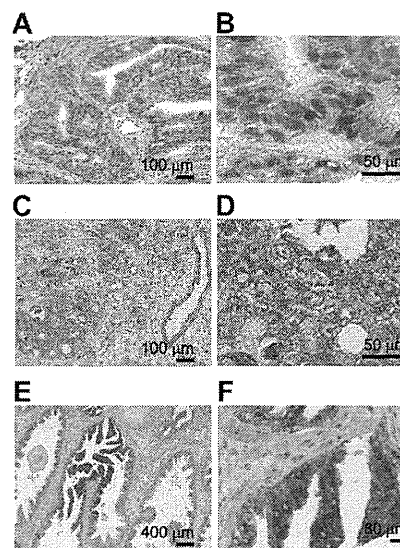


Figure 3. Immunohistochemical staining analysis for pro-neuropeptide Y (NPY). The NPY immunostaining patterns of prostate cancer at Gleason Score 3 + 4 (A) and 4 + 5 (C), respectively. Strong positive immunostaining was observed in the cytoplasm of prostate cancer cells, whereas a weak immunopositivity was found in noncancerous prostate epithelium. (E) Cytoplasmic immunoreactivity with anti-NPY antibody of the high-grade PIN precursor cells. (B,D,F) Five-fold magnified view of A, C, and E, respectively. Scale bars are shown at the right bottom of panels.

are shown. Immunoreactivity with anti-NPY antibody was observed in prostate cancer tissues, exhibiting strong positive immunostaining in the cytoplasm of prostate cancer cells, whereas a weak immunopositivity was found in noncancerous prostate epithelium. Figure 3C,F further indicated that the cytoplasmic immunoreactivity with anti-NPY antibody could be also observed in PIN precursor cells. These results suggested that NPY expression was significantly upregulated at neoplastic lesions in prostate tissue. The staining intensity of NPY seemed to be varied according to Gleason grade of tumor cells within a field.

Table 2. List of Identified Biomarker Candidates for Prostate Cancer

				detection frequency in absent/present search	
UniProt accession number	UniProt ID	protein name	sequence[a]	control[b]	PCa[c]
P01303	NPY_HUMAN	pro-neuropeptide Y	S^{68}SPETLISDLLMR^{80}	0	11
P02751	FINC_HUMAN	fibronectin	Q^{1041}YNVGPSVSK^{1050}	1	11
Q16134	ETFD_HUMAN	electron transfer flavoprotein-ubiquinone oxidoreductase	G^{213}IATNDVGIQK^{223}	1	11

[a]Numbers in the sequences indicate the amino acid numbers. [b]Number of cases detected among healthy controls or BPH patients. [c]Number of cases detected among prostate cancer patients.

39450185

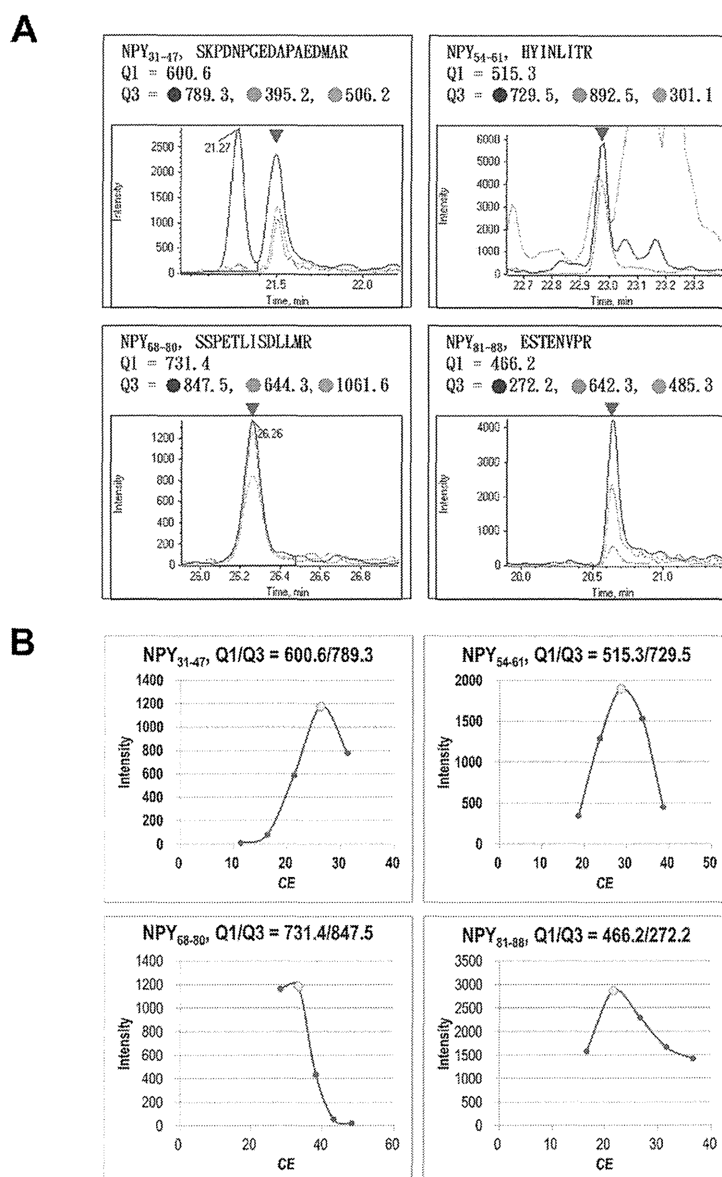


Figure 4. Optimization of MRM conditions for NPY used in the biomarker validation experiment. (A) Selection of the appropriate MRM transition for further quantification analysis. Three transitions were tested for each of four NPY-derived peptides identified in the biomarker screening step. The numbers on the right side of Q1 or Q3 indicate the target setting of m/z in the quadrupole 1 or 3, respectively. (B) Transition showing the highest quality (S/N) of MRM chromatogram was measured with five distinct conditions of collision energy (CE, eV). The CE setting providing the highest intensity of chromatogram peak is indicated by yellow circle.

MRM-Based Replication Analysis for NPY

To evaluate the variability associated with different hospitals or analytical methods, we conducted further validation experiment for NPY using an independent sample set and another mass spectrometric quantification method MRM. Here we replaced healthy controls from Kochi University hospital with ones from JFCR and added prostate disease samples from Kyoto University and Iwate Medical University (Table 1). The 110 QUEST-MS-purified samples were then analyzed in a 4000 QTRAP mass spectrometer by targeting four distinct peptides derived from NPY. Prior to MRM measurement of 110 validation samples, we first searched the optimum transition and CE out of three transitions for each four NPY peptide, which could provide the highest intensity of MRM chromatogram peaks (Figure 4A,B). From the analysis of case-1 in Supplementary Figure S4 in the Supporting Information, the transition Q1/Q3 = 466.2/272.2 corresponding to NPY_{81-88} peptide at CE = 17 eV was found to be the most sensitive reporter ion pair for NPY. Because only NPY_{81-88} transition had enough sensitivity to detect NPY with even normal and BPH control levels, the following NPY quantification analysis was performed by this transition. Additionally, 10 fmol of digested BSA was spiked to each sample and used as the internal control to normalize the interanalytical variability. The quantitative reproducibility of MRM-based NPY measurement

dx.doi.org/10.1021/pr400547s | J. Proteome Res. 2013, 12, 4497−4506

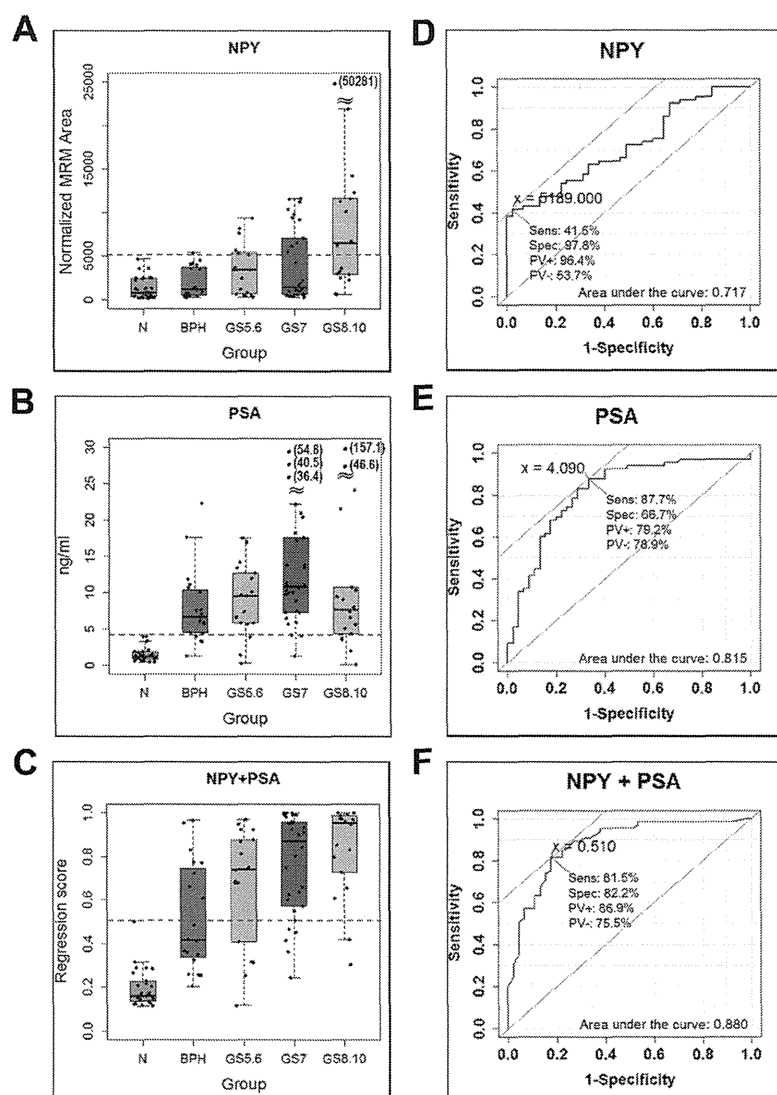


Figure 5. Results of MRM-based validation experiment for NPY using 110 plasma samples. The quantified concentration of NPY by MRM or PSA by ELISA is shown by the box plot (A) or (B), respectively. The combination biomarker of NPY and PSA was constructed by the modified logistic regression method and utilized for the classification of five groups (C). The red dashed lines indicate the thresholds between control group (N + BPH) and prostate cancer group, calculated by the ROC curve analysis in panels D−F. The outliers are displayed separately in the upper area of plots, accompanying each concentration. N: healthy controls, BPH: benign prostate hypertrophy, GS: Gleason score. ROC curves were constructed with plasma concentration of NPY (D), PSA (E), or the regression score of combination marker NPY + PSA (F) by use of 110 validation plasma samples. The point of tangency indicated by "x =" shows the optimum threshold value provided by each biomarker. Sens.: sensitivity, Spec.: specificity, PV+: positive predictive values, PV−: negative predictive values.

was then evaluated by analyzing three plasma samples at six distinct time points in triplicated manner. Supplementary Figure S4 in the Supporting Information showed that observed coefficient of variation (CV) was consistently smaller than 3.5%. Then, by use of optimized MRM conditions above, plasma NPY was measured by relative quantification from 110 cases and displayed on a box plot in Figure 5A, demonstrating much higher diagnostic specificity between BPH patients and Gleason score 5 to 6 prostate cancer patients (p = 0.0396, t test) in comparison with PSA test (p = 0.2746, t test) (Figure 5B). This observation successfully confirmed the prostate-cancer-specific detection of NPY in the absent/present search analysis at the biomarker screening phase. Despite the high specificity, because the sensitivity of NPY appeared not to be sufficient enough for a singly used biomarker, we intended to establish an integrative diagnostic model of NPY and PSA by use of modified logistic regression method (Figure 5C). The logistic regression scores were calculated so that they could provide the maximum area under the curve (AUC) of ROC curve as previously described.[20] Figure 5D−F showed the comparison of three ROC curves generated with NPY (area of MRM chromatogram), PSA (values measured by ELISA), or NPY + PSA (logistic regression scores) of 110 plasma samples, respectively. Whereas the sensitivity and specificity of PSA in distinguishing 65 prostate cancer patients from 45 controls (25 healthy controls and 20 BPH patients) were 87.7 and 66.7%,

respectively, those of the NPY-PSA combination biomarker reached 81.5 and 82.2%, respectively. This result clearly suggested that NPY improved the poor specificity of PSA test without significant loss of its sensitivity, which would lead to reduction of overdiagnosis and unnecessary biopsies, eventually leading to alleviation of both physical and mental stress for men.

DISCUSSION

For the purpose of identifying blood biomarkers for early diagnosis of cancers, comprehensive detection of submicrogram/milliliter proteins should be fundamental because the amount of potential tumor marker proteins secreted from small malignant tissues would be very subtle, even allowing for relatively rapid accumulation of particular proteins which show long half-life in blood, such as CEA (7–14 days),[21] AFP (6.2 days),[22] or CA125 (6 days).[23] To achieve such high sensitivity in plasma proteomics, we focused on <20 kDa LMW subproteome of plasma because most of interfering abundant proteins were found in larger molecular weight area of plasma proteome. The QUEST-MS technology we developed in the present study enabled us not only to enrich the <20 kDa fractions rapidly but also to identify a number of plasma minor components. These facts strongly assured that the sensitivity provided by QUEST-MS technology was high enough to discover very low concentration of biomarker proteins derived from early diseases utilizing any mass spectrometric approaches. This technology can also enhance the detection level of various protein assays, such as ELISA, Western blotting, or HPLC-based analysis targeting small proteins. The considerable benefits of QUEST-MS technology in the clinical application were reproducibility and quickness of experimental procedures. Unlike size exclusion chromatography or ultrafiltration methods, all processes can be executed rapidly by automatic pipetting robots equipped with 96-well syringe pumps (Supplementary Figure S1 in the Supporting Information).

In this report, we applied QUEST-MS technology to the plasma biomarker discovery for prostate cancer by analyzing 116 samples. The whole analytical steps were accomplished within a couple of weeks, involving 1 day of QUEST-MS purification and 12 days of individual LC/MS/MS analyses. The subsequent absent/present search was successful in extracting 189 candidate biomarker peptides showing prostate cancer-specific detection patterns; however, only three of them were sequenced by Sequest database search analysis despite identification of 1126 proteins in total. This poor identification rate might be attributed to quite low concentrations of 189 candidate peaks. Actually, the S/N ratio of mass spectrum peaks for these peptides was only 3–10, resulting in missing MS/MS acquisitions or insufficient quality of MS/MS spectra for database search. As described in the previous paragraph, it is no wonder that early diagnosis biomarkers would be found in the lowest intensity area of mass spectrometric data and the sequence identifications were not always straightforward. Therefore, MudPIT (multidimensional protein identification technology) approaches employing more intensive prefractionation of analytes should be necessary to enforce the identification efficacy, even though the detection in the survey scans has already been achieved.

The new prostate cancer biomarker NPY on which we focused in the present study is one of the well-known neurotransmitter polypeptide composed of 36 amino acids as the activated from in blood. The prepro-NPY (97 amino acids) is synthesized from NPY gene and secreted into extracellular region after removal of the signal peptide NPY_{1-28} (pro-NPY). In the next stage, the C-flanking peptide of NPY (CPON, NPY_{68-97}) is cleaved from pro-NPY by proconverting enzymes, resulting in NPY_{29-67}. Following further cleavage at the C-terminus of Gly^{65} by carboxypeptidase-like enzyme, the Gly^{65} was finally truncated by peptidylglycine-amidating monooxygenase to generate mature NPY_{29-64}. In our biomarker screening data, only NPY_{68-80} tryptic digest was identified, which demonstrated the detection profile as follows (healthy controls + BPH)/(prostate cancer patients) = 0/11. However the MRM-based validation of four distinct NPY tryptic digests, including NPY_{31-47}, NPY_{54-61}, NPY_{68-80}, and NPY_{81-88}, denoted the same tendency of plasma concentration among 110 individuals. These facts could provide two important insights concerning the limit of detection and the dynamic feature of NPY in prostate cancer patients. At first, all of NPY_{31-47}, NPY_{54-61}, and NPY_{81-88} peptides were not involved in 189 biomarker candidates because they were masked by coeluted other peptides and detected only infrequently. Second, the isoform upregulated in prostate cancer patients' plasma was the inactive form pro-NPY (NPY_{29-97}) whose C-terminus had not been removed.

Although major physiological functions of active NPY_{29-64} are known to be strong appetite stimulant,[24] vasoconstrictive effect,[25] and regulation of stress behaviors,[26] the association between NPY level and cancer progression remains controversial. At least the NPY seems to be involved in the development of specific tumors, including neural crest-derived tumors and breast and prostate cancers by facilitating proliferation, invasion, metastasis, and angiogenesis.[27] Importantly, it was reported that overexpression of NPY in prostate cancer tissues had significant correlation with poor prognosis of patients.[28] These reports could support the evidence that a particular group of prostate cancer patients exhibit an aberrantly high level of plasma NPY even at very early grade of cancer such as Gleason score 5 or 6. In fact, the addition of NPY values to PSA test using logistic regression model significantly improved the diagnostic specificity between BPH patients and Gleason score 5 to6 prostate cancer patients (Figure 5).

Recently, two innovative gene diagnostics for prostate cancer, urine PCA3 and TMPRSS2:ERG, have been under clinical reviews. Large-scale trials of urine PCA3 test in the United States (n = 466) and Japan (n = 633) yielded 77.5% sensitivity and 57.1% specificity (PCA3 score threshold = 25)[29] or 66.5% sensitivity and 71.6% specificity (PCA3 score threshold = 35),[30] respectively. TMPRSS2:ERG fusion gene was found in 41 or 43% of primary prostate cancer tissues (n = 59) or castration-resistant prostate cancer tissues (n = 82), respectively.[31] A multivariate regression analysis combining expression levels of PCA3, GOLPH2, SPINK1, and TMPRSS2:ERG in urine cells outperformed serum PSA, resulting in 65.9% sensitivity and 76.0% specificity (n = 234).[32] Even in comparison with these gene diagnostics, our combinational biomarker NPY+PSA illustrated better performance to distinguish prostate cancer from noncancer population (81.5% sensitivity and 82.2% specificity).

CONCLUSIONS

In summary, we described the development of versatile LMW-focusing proteomics technology QUEST-MS and the application of this method to plasma biomarker discovery for prostate

cancer. QUEST-MS technology provides an adequate reduction in sample complexity to improve detection of low abundance proteins. Although here we utilized Oasis HLB μElution plates for QUEST-MS purification, other reversed-phase chromatographic materials with different functional moieties or pore sizes would expand the variety of purified molecules. In future applications, the automated QUEST-MS procedure will drastically facilitate sensitive and high-throughput analysis of small bioactive polypeptides, such as hormones, neurotransmitters, or cytokines, from any body fluids including serum, cerebrospinal fluid, or urine.

■ ASSOCIATED CONTENT

Supporting Information

Automatic 96-well syringe robot used for QUEST-MS purification, workflow for the label-free quantification analysis on the Expressionist RefinerMS module, molecular weight distribution of identified proteins from QUEST-MS-purified plasma samples, evaluation of variability for MRM-based plasma NPY measurements, and a list of 1126 nonredundant proteins identified from biomarker screening phase. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +81-3-6409-2306. Fax: +81-3-5449-5785. E-mail: k-ueda@riken.jp.

Notes

The authors declare the following competing financial interest(s): A.T. and T.A.S. are employees of Shimadzu Corporation. M.N. is employee of Toppan Printing Co., Ltd.

■ ACKNOWLEDGMENTS

This work was supported by funding to Koji Ueda. from Grant-in-Aid for Young Scientists (A), The Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. Shimadzu Corporation and Toppan Printing Co., Ltd. also financially supported the present study.

■ ABBREVIATIONS

QUEST-MS, quick enrichment of small targets for mass spectrometry; BPH, benign prostate hypertrophy; PSA, prostate-specific antigen; NPY, neuropeptide-Y; PIN, prostatic intraepithelial neoplasia; MRM, multiple reaction monitoring; LMW, low-molecular-weight

■ REFERENCES

(1) Ablin, R. J.; Pfeiffer, L.; Gonder, M. J.; Soanes, W. A. Precipitating antibody in the sera of patients treated cryosurgically for carcinoma of the prostate. *Exp. Med. Surg.* **1969**, *27* (4), 406–10.

(2) Catalona, W. J.; Smith, D. S.; Ratliff, T. L.; Dodds, K. M.; Coplen, D. E.; Yuan, J. J.; Petros, J. A.; Andriole, G. L. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N. Engl. J. Med.* **1991**, *324* (17), 1156–61.

(3) Chou, R.; Croswell, J. M.; Dana, T.; Bougatsos, C.; Blazina, I.; Fu, R.; Gleitsmann, K.; Koenig, H. C.; Lam, C.; Maltz, A.; Rugge, J. B.; Lin, K. Screening for prostate cancer: a review of the evidence for the U.S. Preventive Services Task Force. *Ann. Intern. Med.* **2011**, *155* (11), 762–71.

(4) Brett, A. S.; Ablin, R. J. Prostate-cancer screening–what the U.S. Preventive Services Task Force left out. *N. Engl. J. Med.* **2011**, *365* (21), 1949–51.

(5) Kurogochi, M.; Matsushista, T.; Amano, M.; Furukawa, J.; Shinohara, Y.; Aoshima, M.; Nishimura, S. Sialic acid-focused quantitative mouse serum glycoproteomics by multiple reaction monitoring assay. *Mol. Cell. Proteomics* **2010**, *9* (11), 2354–68.

(6) Ueda, K.; Takami, S.; Saichi, N.; Daigo, Y.; Ishikawa, N.; Kohno, N.; Katsumata, M.; Yamane, A.; Ota, M.; Sato, T. A.; Nakamura, Y.; Nakagawa, H. Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Mol. Cell. Proteomics* **2010**, *9* (9), 1819–28.

(7) Villanueva, J.; Martorella, A. J.; Lawlor, K.; Philip, J.; Fleisher, M.; Robbins, R. J.; Tempst, P. Serum peptidome patterns that distinguish metastatic thyroid carcinoma from cancer-free controls are unbiased by gender and age. *Mol. Cell. Proteomics* **2006**, *5* (10), 1840–52.

(8) Ueda, K.; Saichi, N.; Takami, S.; Kang, D.; Toyama, A.; Daigo, Y.; Ishikawa, N.; Kohno, N.; Tamura, K.; Shuin, T.; Nakayama, M.; Sato, T. A.; Nakamura, Y.; Nakagawa, H. A comprehensive peptidome profiling technology for the identification of early detection biomarkers for lung adenocarcinoma. *PloS One* **2011**, *6* (4), e18567.

(9) Shen, Y.; Liu, T.; Tolic, N.; Petritis, B. O.; Zhao, R.; Moore, R. J.; Purvine, S. O.; Camp, D. G.; Smith, R. D. Strategy for degradomic-peptidomic analysis of human blood plasma. *J. Proteome Res.* **2010**, *9* (5), 2339–46.

(10) Liu, T.; Qian, W. J.; Mottaz, H. M.; Gritsenko, M. A.; Norbeck, A. D.; Moore, R. J.; Purvine, S. O.; Camp, D. G., 2nd; Smith, R. D. Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol. Cell. Proteomics* **2006**, *5* (11), 2167–74.

(11) Mouton-Barbosa, E.; Roux-Dalvai, F.; Bouyssie, D.; Berger, F.; Schmidt, E.; Righetti, P. G.; Guerrier, L.; Boschetti, E.; Burlet-Schiltz, O.; Monsarrat, B.; Gonzalez de Peredo, A. In-depth exploration of cerebrospinal fluid by combining peptide ligand library treatment and label-free protein quantification. *Mol. Cell. Proteomics* **2010**, *9* (5), 1006–21.

(12) Capriotti, A. L.; Caruso, G.; Cavaliere, C.; Piovesana, S.; Samperi, R.; Lagana, A. Comparison of three different enrichment strategies for serum low molecular weight protein identification using shotgun proteomics approach. *Anal. Chim. Acta* **2012**, *740*, 58–65.

(13) Hood, B. L.; Lucas, D. A.; Kim, G.; Chan, K. C.; Blonder, J.; Issaq, H. J.; Veenstra, T. D.; Conrads, T. P.; Pollet, I.; Karsan, A. Quantitative analysis of the low molecular weight serum proteome using 18O stable isotope labeling in a lung tumor xenograft mouse model. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (8), 1221–30.

(14) Lassout, O.; Pastor, C. M.; Fetaud-Lapierre, V.; Hochstrasser, D. F.; Frossard, J. L.; Lescuyer, P. Analysis of the pancreatic low molecular weight proteome in an animal model of acute pancreatitis. *J. Proteome Res.* **2010**, *9* (9), 4535–44.

(15) Wu, J.; An, Y.; Pu, H.; Shan, Y.; Ren, X.; An, M.; Wang, Q.; Wei, S.; Ji, J. Enrichment of serum low-molecular-weight proteins using C18 absorbent under urea/dithiothreitol denatured environment. *Anal. Biochem.* **2010**, *398* (1), 34–44.

(16) Wang, Y. Y.; Lee, C. T.; Lu, C. L.; Chen, C. Y.; Chang, F. Y.; Lee, S. D.; Doong, M. L.; Wang, P. S. Gastric inhibitory polypeptide appears less important in mediating acid secretion. *Hepato-Gastroenterology* **1999**, *46* (27), 2105–9.

(17) Sehested, A.; Juul, A. A.; Andersson, A. M.; Petersen, J. H.; Jensen, T. K.; Muller, J.; Skakkebaek, N. E. Serum inhibin A and inhibin B in healthy prepubertal, pubertal, and adolescent girls and adult women: relation to age, stage of puberty, menstrual cycle, follicle-stimulating hormone, luteinizing hormone, and estradiol levels. *J. Clin. Endocrinol. Metab.* **2000**, *85* (4), 1634–40.

(18) Di Iorio, R.; Marinoni, E.; Letizia, C.; Villaccio, B.; Alberini, A.; Cosmi, E. V. Adrenomedullin production is increased in normal human pregnancy. *Eur. J. Endocrinol.* **1999**, *140* (3), 201–6.

(19) Su, A. I.; Wiltshire, T.; Batalov, S.; Lapp, H.; Ching, K. A.; Block, D.; Zhang, J.; Soden, R.; Hayakawa, M.; Kreiman, G.; Cooke, M. P.; Walker, J. R.; Hogenesch, J. B. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (16), 6062–7.

(20) Pepe, M. S.; Cai, T.; Longton, G. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* **2006**, *62* (1), 221−9.

(21) Choi, J. S.; Min, J. S. Significance of postoperative serum level of carcinoembryonic antigen (CEA) and actual half life of CEA in colorectal cancer patients. *Yonsei Med. J.* **1997**, *38* (1), 1−7.

(22) Gerl, A.; Lamerz, R.; Clemm, C.; Mann, K.; Hartenstein, R.; Wilmanns, W. Does serum tumor marker half-life complement pretreatment risk stratification in metastatic nonseminomatous germ cell tumors? *Clin. Cancer Res.* **1996**, 2 (9), 1565−70.

(23) Meyer, T.; Rustin, G. J. Role of tumour markers in monitoring epithelial ovarian cancer. *Br. J. Cancer* **2000**, *82* (9), 1535−8.

(24) Zhang, L.; Bijker, M. S.; Herzog, H. The neuropeptide Y system: pathophysiological and therapeutic implications in obesity and cancer. *Pharmacol. Ther.* **2011**, *131* (1), 91−113.

(25) Hirsch, D.; Zukowska, Z. NPY and stress 30 years later: the peripheral view. *Cell. Mol. Neurobiol.* **2012**, *32* (5), 645−59.

(26) Sah, R.; Geracioti, T. D. Neuropeptide Y and posttraumatic stress disorder. *Mol. Psychiatry* **2013**, *18* (6), 646−55.

(27) Ruscica, M.; Dozio, E.; Motta, M.; Magni, P. Relevance of the neuropeptide Y system in the biology of cancer progression. *Curr. Top. Med. Chem.* **2007**, *7* (17), 1682−91.

(28) Rasiah, K. K.; Kench, J. G.; Gardiner-Garden, M.; Biankin, A. V.; Golovsky, D.; Brenner, P. C.; Kooner, R.; O'Neill, G, F.; Turner, J. J.; Delprado, W.; Lee, C. S.; Brown, D. A.; Breit, S. N.; Grygiel, J. J.; Horvath, L. G.; Stricker, P. D.; Sutherland, R. L.; Henshall, S. M. Aberrant neuropeptide Y and macrophage inhibitory cytokine-1 expression are early events in prostate cancer development and are associated with poor prognosis. *Cancer Epidemiol., Biomarkers Prev.* **2006**, *15* (4), 711−6.

(29) Gittelman, M.; Hertzman, B.; Bailen, J.; Williams, T.; Koziol, I.; Henderson, R. J.; Efros, M.; Bidair, M.; Ward, J. F. PCA3 molecular urine test as a predictor of repeat prostate biopsy outcome in men with previous negative biopsies: A prospective multicenter clinical study. *J. Urol.* **2013**, *190* (1), 64−69.

(30) Ochiai, A.; Okihara, K.; Kamoi, K.; Oikawa, T.; Shimazui, T.; Murayama, S. I.; Tomita, K.; Umekawa, T.; Uemura, H.; Miki, T. Clinical utility of the prostate cancer gene 3 (PCA3) urine assay in Japanese men undergoing prostate biopsy. *BJU Int.* **2013**, *111* (6), 928−33.

(31) Qu, X.; Randhawa, G.; Friedman, C.; O'Hara-Larrivee, S.; Kroeger, K.; Dumpit, R.; True, L.; Vakar-Lopez, F.; Porter, C.; Vessella, R.; Nelson, P.; Fang, M. A novel four-color fluorescence in situ hybridization assay for the detection of TMPRSS2 and ERG rearrangements in prostate cancer. *Cancer Genet.* **2013**, *206* (1−2), 1−11.

(32) Laxman, B.; Morris, D. S.; Yu, J.; Siddiqui, J.; Cao, J.; Mehra, R.; Lonigro, R. J.; Tsodikov, A.; Wei, J. T.; Tomlins, S. A.; Chinnaiyan, A. M. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. *Cancer Res.* **2008**, *68* (3), 645−9.

Review

Glycoproteomic strategies: From discovery to clinical application of cancer carbohydrate biomarkers

Koji Ueda

Laboratory for Biomarker Development, Center for Genomic Medicine, RIKEN, Minato-ku, Tokyo, Japan

Carbohydrate antigens are the most frequently and traditionally used biomarkers for cancer, such as CA19–9, CA125, DUPAN-II, AFP-L3, and many others. The diagnostic potential of them was simply based on the cancer-specific alterations of glycan structures on particular glycoproteins in serum/plasma. In spite of the facts that glycosylation disorders are feasible for cancer biomarkers and glycomic analysis technologies to explore them have been rapidly developed, it remains difficult to sensitively screen glycan structure changes on cancer-associated glycoproteins from clinical specimens. Moreover, a lot of additional issues should be appropriately addressed for the clinical application of newly identified glycosylation biomarkers, including analytical throughput, quantitative confirmation of structural changes, and biological explanation for the alterations. In the last decade, significant improvement of mass spectrometric techniques is being made in the aspects of both hardware spec and preanalytical purification procedures for glycoprotein analysis. Here we review potential approaches to perform comprehensive analysis of glycoproteomic biomarker screening from serum/plasma and to realize high-throughput validation of site-specific oligosaccharide variations. The power and problems of mass spectrometric applications on the clinical use of carbohydrate biomarkers are also discussed in this review.

Received: November 29, 2012
Accepted: December 27, 2012

Keywords:
Biomarker / Cancer / Glycoproteomics / Mass spectrometry / Serum

1 Introduction

In order to survey tumor-derived biomarkers in serum, not only glycoproteomics but all other focusing technologies are essential, which effectively reduce the complexity of samples by removing unnecessary major serum proteins [1, 2]. Such focused proteomics approaches would include peptidomics [3–6], degradomics [7–11], immunodepletion of abundant proteins [12–16], ProteoMiner purification [17–22], and rare amino acid capturing (e.g. iCAT tag [23–28] for cysteine or 2-nitrobenzenesulfenyl tag for tryptophan [29–34]). Since concentration of biomarker molecules released from early-stage cancer tissues (generally less than 20 mm diameter) are extremely low, appropriate focused proteomics technologies can greatly help to detect such sub-μg/mL proteins in sera. In particular, glycosylation-focused proteomics could contribute to both improving analytical depth and interpreting roles of carcinogenesis-associated glycosylations, which might lead to further understanding molecular mechanisms of malignant transformation or metastasis. More importantly, particular alterations of glycan structures on secreted glycoproteins specifically reflect the site of original organs or cell types. This aspect of glycosylation biomarkers enforces the specificity of diagnosis in the cases distinguishing cancer from benign diseases (e.g. lung cancer from chronic obstructive pulmonary disease) or defining origin of cancer. Despite the benefits described above, there have been lots of technical difficulties in the discovery of cancer-associated glycosylation microheterogeneities from clinical specimens as well as clinical application of those. Indeed, the latest series of mass spectrometers can provide much deeper and wider knowledge of proteins compared to those a decade ago. However, it has been realized that mass spectrometric analysis of lectin-purified serum proteins, which is one of the simplest ways of glycoproteomics, is far from sufficient to

Correspondence: Dr. Koji Ueda, Laboratory for Biomarker Development, Center for Genomic Medicine, RIKEN, Institute of Medical Science, General Research Building 6F, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
E-mail: k-ueda@riken.jp
Fax: +81-3-5449-5785

Abbreviations: Erexim, energy-resolved oxonium ion monitoring; **GalNAz,** tetra-aceylated-*N*-azidoacetylgalactosamine; **GpIb,** glycoprotein Ib; **IGEL,** isotopic glycosidase elution and labeling on lectin-column chromatography; **WGA,** wheat germ agglutinin

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

identify glycosylation biomarkers for early cancer diagnosis. Here we discuss about issues and resolutions in the use of lectins by introducing methods for sample preparation prior to lectin chromatography. In addition to lectin-based approaches, which have been most frequently utilized in glycomics, other sophisticated glycocapturing technologies are reviewed. Because the concept required for biomarker discovery phase or preclinical validation phase is fundamentally different, the adequate methodologies for each are separately described.

2 Glycoproteomics for biomarker screening

2.1 Lectins or chemicals

Nowadays, state-of-the-art glycoproteomic technologies have been developed to enrich glycoproteins or glycopeptides from crude serum samples. These technologies are separated into lectin-based methods and chemical-based methods, in principle. Which is better for the purpose of carbohydrate-targeting tumor marker discovery? If we intend to identify glycan structure changes as tumor markers, rather than concentration of core proteins, lectin is the only enrichment tool recognizing specific oligosaccharide linkages, excepting sialic acid specific chemistry reverse glycoblotting [35, 36]. The chemical enrichment of glycopeptides, such as using hydrazide chemistry [37–42], boronic acid [43–45], or hydrophilic interactions [46–48], certainly exhibits rigid interaction with glycan moieties, whereas most of lectin–glycan interactions are fragile [49]. However, all of the chemical approaches above are based on covalent or affinity bond with rich hydroxyl groups on oligosaccharides, resulting in comprehensive and structure-unspecific capture of glycopeptides. Therefore, the chemical route is inadvisable for glycan structure-targeting biomarker discovery. From the view of such features, we would like to focus on lectin-based glycopeptide enrichment methods in the following sections.

2.2 Using lectin column chromatography for glycoproteomics

Assuming that the eluate of lectin column chromatography would be analyzed in LC/MS/MS, it must be a critical issue whether we load proteins or digested peptides to lectin columns. When undigested serum proteins are purified with lectin column chromatography and eluted by hapten sugars, targeted glycoproteins would be eluted with a lot of nonspecific proteins and high concentration of hapten sugars, which could not be appropriate for mass spectrometric analysis (Fig. 1A). The large amount of nonglycosylated protein elution is mainly caused by the limitation of solvent used in lectin column chromatography, with which the use of detergents, high salts, and organic solvents are not compatible. Furthermore, since most of serum proteins form complexes, coelution of intact binding proteins is inevitable. More importantly, if enriched glycoproteins would have multiple glycosylation sites, it is hardly distinguishable which glycosylation sites might be associated with state of cancer.

On the other hand, when digested peptides are loaded to lectin columns, the purification efficiency of glycopeptides is relatively high due to the elimination of protein–protein interaction effect in the samples (Fig. 1B). However, elution of glycopeptides using hapten sugars still results in significant contamination of 200–1000 mM sugars in the eluate, which is not adequate for direct injection to LC/MS/MS. In addition to this fact, mass spectrometric analysis of eluted glycopeptides is also inefficient for comprehensive studies because automated protein identification by database search is impossible for glycopeptides, for which additional deglycosylation steps would be finally required.

These characteristics of lectins on glycoproteomics have led to recent development of glycopeptidase elution technologies [46, 50]. After binding glycopeptides on lectin columns, PNGase-F elution in volatile salt buffers allows highly specific elution of only originally glycosylated peptides. After lyophilizing eluate, the final product consists of completely deglycosylated peptides with no contaminant salts, which could be directly injected into LC/MS instruments and subjected to usual database search analysis on Mascot or Sequest software. Using this type of enzymatic elution procedure on click maltose HILIC beads, Zhu et al. successfully identified 92, 178, and 221 unique *N*-glycosylation sites from 10 nL, 100 nL, and 1 μL of human serum, respectively [46]. They effectively excluded desalting, buffer exchanging, and lyophilization steps to finish all preanalytical procedures by spin columns within 1.5 h.

Regarding peptide sequencing of enzymatically deglycosylated peptides on database search analysis, *N*-glycosylation sites are recognized as aspartic acid residues converted from asparagine residues by PNGase-F (Fig. 1C). However, chemically identical conversion may artificially occur on asparagine residues known as deamidation of asparagine. To eliminate the false-positive identification of *N*-glycosylation sites by deamidation, we can utilize the PNGase-F reaction in heavy water ($H_2{}^{18}O$) [50–57]. The incorporation of ^{18}O into glycosylated asparagine residues by PNGase-F induces generation of 3 Da increased asparagine residues, providing N-glycosylation site-specific stable isotope tags on peptides (Fig. 1C). Our team recently integrated on-column PNGase-F elution with the ^{18}O stable isotope labeling method and reported as an effective glycoproteomic biomarker screening technology, named isotopic glycosidase elution and labeling on lectin-column chromatography (IGEL) [50].

Hence, issues on biochemical properties of lectins for glycoproteomics can be overcome by employing both protease digestion before lectin column purification and on-column glycopeptidase elution of peptides.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

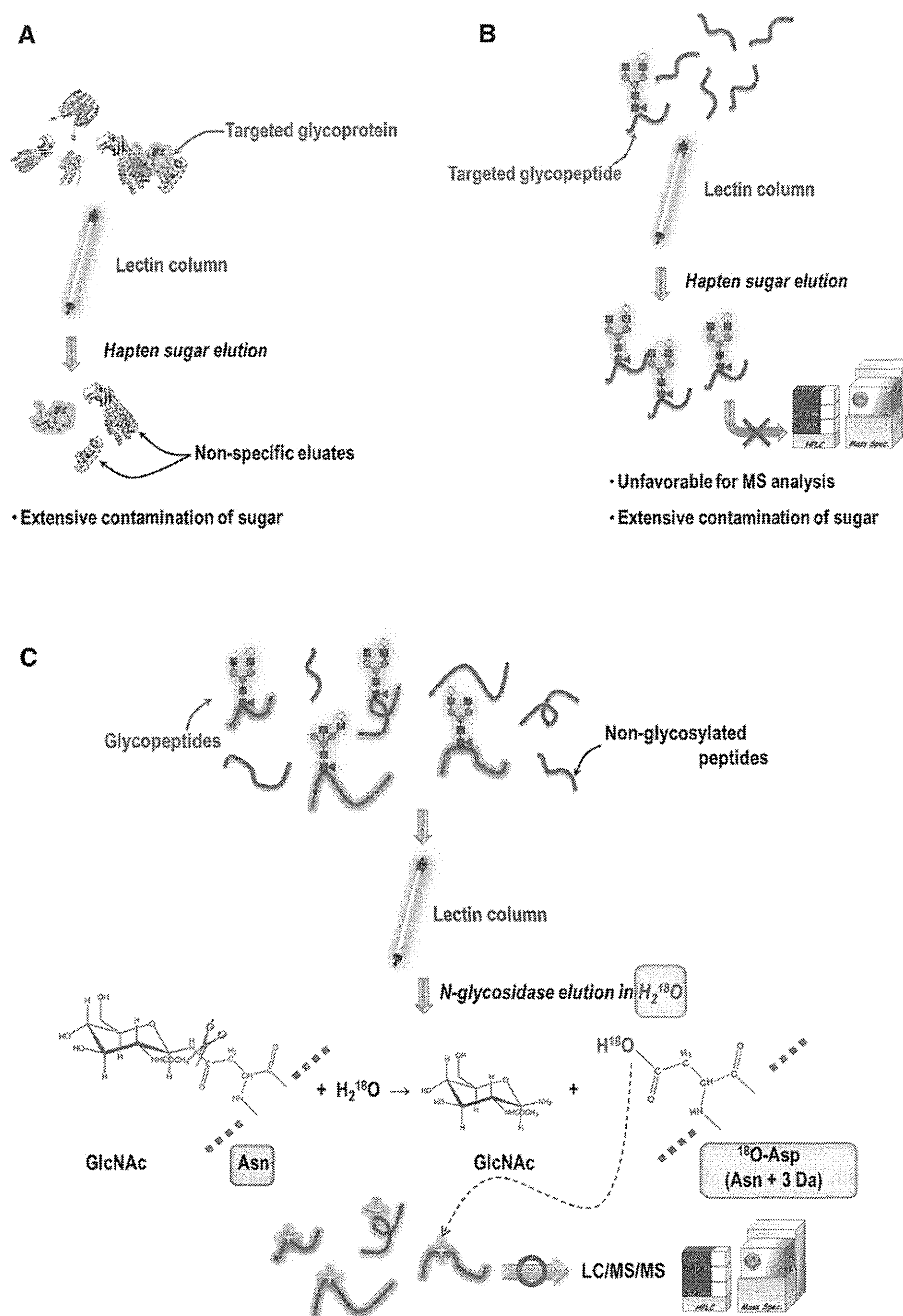


Figure 1. Difficulties and resolution for use of lectin column chromatography in glycoproteomics. (A) Since detergent, high salt, and organic solvent are unusable, lectin column purification using undigested protein samples results in significant contamination of nonspecific proteins due to protein–protein interactions. High concentration of eluting salt is also involved in final product. (B) When starting with digested peptide mixtures, glycopeptide enrichment yield is relatively high. However, direct analysis of eluted glycopeptides is difficult because of low ionization efficiency and impossibility of automated database search. (C) When captured glycopeptides are eluted with *N*-glycosidase, highly specific elution of originally glycosylated peptide could be achieved. If this reaction is combined with ^{18}O-labeling method, direct LC/MS/MS analysis can identify glycosylation sites specifically as 3 Da increased asparagine residues.

2.3 Pre-enrichment of glycopeptides prior to lectin purification

Lectin columns are convenient and widely used enrichment tools for glycoproteins or glycopeptides. As contrasted with the benefits, enrichment ratio itself is not necessarily high enough. When we purified tryptic digest of crude serum by various types of nine different lectin columns according to IGEL method described above, the enrichment efficiency (number of glycopeptide identification / total peptide identification × 100) was only 20–45%. Such insufficient enrichment efficiency was mainly caused by weak lectin–oligosaccharide affinities and the fact that abundance of nonglycosylated peptides in tryptic digest of crude serum samples was absolutely higher than that of glycosylated peptides. This aspect emphasizes the importance of adequate pre-enrichment techniques for pan-glycosylated peptides, such as cellulose column [58–60], graphite carbon column [61–65], and other

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

hydrophilic affinity resins. Sergei et al. constructed C_{18}-cellulose mix-mode column chromatography and achieved high-yield extraction of *N*- and *O*-glycosylated peptides from mixture of ten standard proteins [59]. Lam et al. applied an online combination of RP/RP and porous graphite carbon LC to the comprehensive analysis of ConA lectin-purified human serum samples and identified 134 *N*-glycosylated serum proteins, 151 possible *N*-glycosylation sites, and more than 40 possible *N*-glycan structures [65]. The CL-4B Sepharose-based hydrophilic extraction of glycopeptides is also one of the most popular techniques in glycoproteomics. Selman et al. loaded 5 μL of CL-4B Sepharose beads into 96-well format plate and enriched human IgG-derived glycopeptides for MALDI-FT-ICR-MS [66]. This type of multiplexed purification system is especially suitable for large-scale biomarker screening assays. Actually, we also applied CL-4B Sepharose beads pre-enrichment procedure before lectin column purification and obtained finally around 90% glycopeptide enrichment ratio [50]. Thus pre-enrichment of total glycopeptides from complex peptide mixtures can drastically improve the glycopeptide focusing efficiency by subsequent lectin column chromatography.

2.4 Selection of lectins for glycoproteomics

So far hundreds of lectins have been isolated from plants, microbes, or animals and most of them are commercially available. The selection of lectins is a critical step for precise and comprehensive profiling of cancerous glycan disorders. In the glycoproteomic studies, both high specificity for glycan structures and high affinity to capture glycopeptides are needed. The glycopeptide enrichment ratio acquired from IGEL purification experiments using human serum and nine distinct lectins (LCA, SNA-I, SNA-II, UEA-I, wheat germ agglutinin (WGA), LPA, ConA, and SSA) was shown in Table 1. Concerning specificity of lectins, LCA, SNA-I, ConA, and SSA demonstrated over 80% glycopeptide enrichment rate, suggesting that ligand specificity of these four lectins would be sufficient. Meanwhile, when looking into numbers of glycopeptide identification, which indirectly reflected the lectin–glycan affinity of each lectin, LCA and SNA-I lectins showed much less glycopeptide recovery rate compared to ConA or SSA. Additionally, the remaining five lectins (Lotus, SNA-II, UEA-I, WGA, and LPA) were scarcely able to capture peptides. At least in our binding condition (100 mM ammonium bicarbonate, 5% ACN, 1 mM calcium chloride, 1 mM manganese chloride), ConA and SSA lectin columns could be considered as appropriate materials to be used for specific and comprehensive profiling of human serum glycoproteome. Individual optimization would be required when analyzing other biological samples or using different condition of binding buffers. In order to cover a larger number of glycan structure changes, it is fundamental to increase options of lectins along with optimum purification protocols.

Table 1. Serum glycopeptide enrichment rate by nine lectin columns

Lectin	Protein IDs[a]	Peptide IDs[b]	IGEL(+)[c]	IGEL (%)[d]
LCA	109	173	153	88.3
Lotus	51	27	10	36.7
SNA-I	110	194	186	95.9
SNA-II	62	57	33	57.6
UEA-I	13	13	6	46.4
WGA	82	78	41	52.4
LPA	26	14	5	35.7
ConA	183	413	388	93.9
SSA	229	519	425	81.9

a) Number of identified proteins.
b) Number of identified peptides.
c) Number of peptides possessing IGEL tags (3 Da increased asparagine residues).
d) Ratio of IGEL-tagged peptides in total peptide identification, indicating glycopeptide enrichment rate.

In recent studies, multilectin affinity chromatography approaches were developed to enhance glycopeptide recovery rate and expand the comprehensiveness of targeted glycan structures. Zeng et al. combined high abundance protein depletion, ConA-Jacalin-WGA multilectin affinity chromatography, IEF separation, and LC-MS analysis and identified breast cancer associated proteins such as thrombospondin-1 and 5, alpha-1B-glycoprotein, serum amyloid P-component, and tenascin-X, which had potentially abnormal glycans [67]. The same group rigorously evaluated the identical lectin mixture in several glycoproteome profiling studies [68–72]. Qiu et al. integrated serial lectin purification by ConA and SNA with d0- or d3-*N*-acetoxysuccinamide stable isotope labeling on α-amino groups. By use of this methodology, they enabled effective enrichment of sialylated glycopeptides and also differential analysis of those [73].

2.5 Quantitative assessment of glycan structure changes

To identify cancer-associated alterations of glycosylation on multiple proteins, establishment of rigorous quantification strategies should be essential, which could stoichiometrically evaluate the changing rate of each glycoform. Comparative quantification results of only enriched glycopeptides are affected by not only glycan structure changes but also concentration of original core proteins itself, indicating that it is hard to determine whether the identified candidates might be glycosylation-targeting biomarkers or protein concentration biomarkers. Therefore, subtraction of protein concentration effects from quantification results of lectin-enriched glycopeptides is necessary (Fig. 2). Recently, we demonstrated a practical example of this concept for the identification of carbohydrate-targeting lung cancer biomarker discovery [50]. Here we acquired relative quantification profiles from both lectin-purified glycopeptides

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

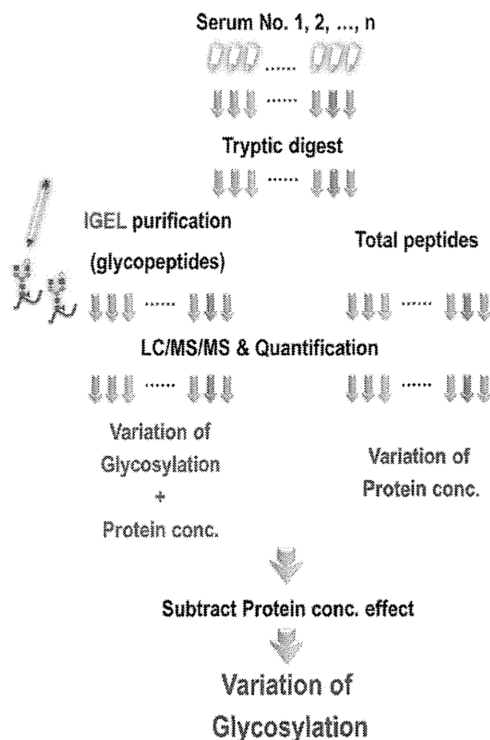


Figure 2. Representative scheme for comprehensive quantification of glycan structure alterations. Following tryptic digestion of serum samples, LC/MS/MS analysis of enriched glycopeptides and total peptides should be performed individually in order to eliminate the effect of core protein concentration variations. Finally, the changing rate of glycans on each peptide could be calculated by subtracting quantification results of total peptides from those of enriched glycopeptides. The representative method for glycopeptide enrichment is IGEL technology illustrated in Fig. 1C.

and pre-enriched nonglycosylated peptides by LC/MS/MS analysis individually. Finally, the site-specific glycoform changes were determined by subtracting core protein concentrations calculated by nonglycosylated peptides from quantification results of each glycopeptide. This approach can illuminate the glycan structure changes on diverse glycosylation sites individually.

2.6 Quantitative glycoproteomic approaches for *O*-glycans

Although tools for comprehensive analysis of *O*-glycosylation are still limited compared with the *N*-glycomics, recent development of sophisticated chemistries have potential to be breakthrough technologies for *O*-glycan biomarker discovery. Hang et al. developed a metabolic labeling approach, which utilized incorporation of tetra-aceylated-*N*-azidoacetylgalactosamine (GalNAz) into the reducing terminus of *O*-glycosylation sites [74, 75]. Following cell culture in the presence of GalNAz for several days, *O*-glycosylated proteins can be specifically collected by alkyne-activated resins (click chemistry) (Fig. 3). Using this technology, Slade et al. identified 267 potentially *O*-glycosylated proteins from the secretome of CHO cells [76]. Furthermore, the Bertozzi's group succeeded to apply GalNAz chemistry to rapid profiling of *O*-linked glycoproteins in living mice [77] and also in vivo imaging of membrane-associated glycans in zebrafish [78]. The other group expanded this technology by using GalNAz, ManNAz, and GlcNAz to discover cell surface differentiation markers on human mesenchymal stem cells [79]. Integration of such metabolic labeling methods for *O*-glycans with mass spectrometric structural analysis may facilitate comprehensive screening of fine *O*-glycan structure alterations in the future.

3 Glycoproteomics for high-throughput biomarker validation

For clinical application of biomarkers, preclinical validation experiments are usually required using independent larger sample set and high-throughput quantification methods. In general, most of protein biomarkers are quantitatively measured by immunoassays in the validation phase, which include sandwich ELISA, AlphaLISA [80–84], and Luminex technology [85–89]. However, it is often extremely difficult to make specific antibodies against both detailed glycan structures and glycosylation sites mainly because of the low immunogenicity of oligosaccharides and structural hindrance of amino acid epitopes by glycans. Thus, alternative glycoproteomic technologies must be established for the replication assays using hundreds of clinical samples, which could detect site-specific glycan structure changes quantitatively with high-throughput manner from complex protein mixtures, such as crude serum/plasma. Here we introduce a couple of technologies used in the glycosylation biomarker validation area.

3.1 Lectin-antibody sandwich ELISA

A sandwich-type ELISA using lectin- and analyte-specific antibody was originally developed by Drouin et al. in 1988 [90]. They intended to establish high-throughput diagnostic assay for Bernard Soulier syndrome, which was a rare autosomal recessive coagulopathy leading to a deficiency of glycoprotein Ib (GpIb), an important clotting regulator as the receptor for von Willebrand factor. In that study they immobilized WGA lectin on 96-well plates to capture GpIb in prepared human platelet proteins and eventually succeeded to detect GpIb by specific monoclonal antibody AN51 quantitatively. They mentioned that this approach was simple, inexpensive, and sensitive way to quantitate glycoproteins for which specific lectins and monoclonal antibodies were available. Tojo et al.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

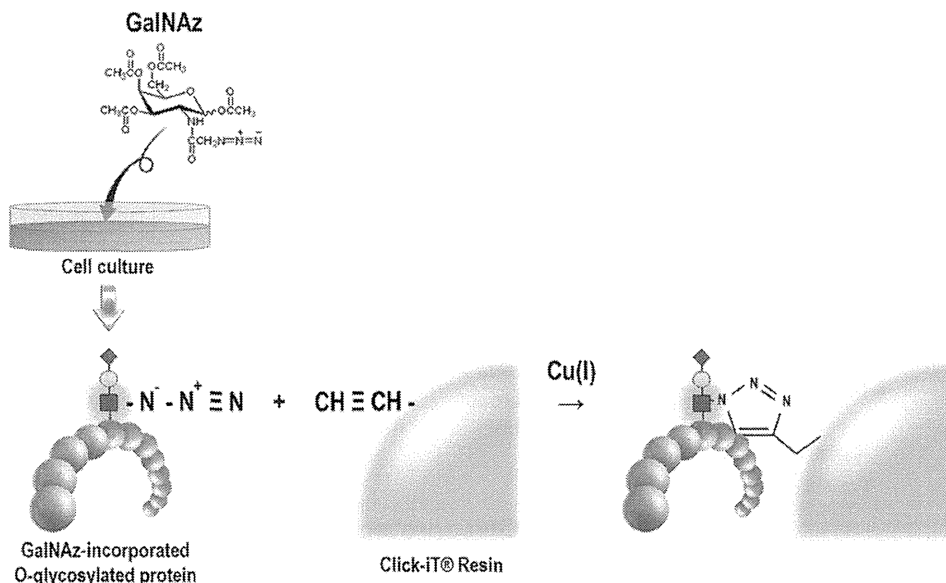


Figure 3. Principle of GalNAz-based *O*-glycosylated protein profiling. Metabolic incorporation of GalNAz residue at the reducing terminus of *O*-glycans enables subsequent capture by alkene reactive resins.

applied the same concept of assay utilizing ConA-immobilized ELISA plates and specific polyclonal antibody to quantitate D-mannans of *Candida albicans* [91]. Since ConA lectin possessed a high binding specificity for the D-mannopyranose unit, the sensitivity and specificity of D-mannan detection were significantly improved compared to previous antibody–antibody ELISA or quantitative precipitin reaction, they reported.

In cancer diagnostic studies, Parker reported the application of lectin/antibody sandwich ELISA assay to the serological diagnosis of pancreatic cancer [92]. His team captured glycoproteins possessing *N*-acetylglucosamie and sialic acid moieties by WGA lectin and detected by CAM17.1 monoclonal antibody specific to a part of mucins. They provided diagnostic assay results from not only retrospective study showing a sensitivity of 78% for pancreatic cancer with a specificity of 76% [93], but also prospective study showing even better sensitivity and specificities (84 and 92%, respectively), suggesting that the assay probably performed better on fresh samples. They also circumstantially summarized the usefulness of CAM 17.1/WGA test for pancreatic cancer diagnosis in *Lancet* journal [94]. Very recently, Miyoshi's group found that fucosylated haptoglobin had a great potential for the detection of pancreatic cancer and prognosis of postoperative colorectal cancer [95, 96]. They coated 96-well plates with antihaptoglobin Fab antibody because IgG had fucosylated oligosaccharides in its Fc portion, and detected fucosylated haptoglobin by biotinylated AAL lectin [95–98]. They measured serum samples from 397 individuals and concluded that the sensitivity and specificity for the diagnosis of pancreatic cancer patients from normal controls was 50 and 91%, respectively [98].

3.2 Energy-resolved oxonium ion monitoring (Erexim) technology

For the purpose of first high-throughput site-specific quantification of glycan structure variations, we recently developed Erexim technology [99]. The oxonium ions are defined as any oxygen cations with three bonds in chemistry, while they are used synonymously with sugar oxonium ions in glycoproteomics, which are produced as oligosaccharide fragment ions in collision cells of mass spectrometers during CID of glycopeptides [100]. Typical oxonium ions (m/z) and the corresponding oligosaccharide components are shown in the right side of Fig. 4. The detection of particular set of oxonium ions in MS/MS spectra of glycopeptides not only represents the existence of glycan modifications but also provides signature of original glycan structures [101]. Furthermore, we found that monitoring the yields of oxonium ions over a wide range of collision energy by use of MRM [102–104] on quadrupole mass spectrometer exhibited a highly glycan structure-unique fragmentation patterns. Indeed, the Erexim curves allowed us to clearly distinguish even structural isomers (same mass with distinct glycan linkage) and separately quantify their contents (Fig. 4). This technology was then applied to the *N*-glycan profiling of three model therapeutic antibody drugs, Herceptin (trastuzumab), Avastin (bevacizumab), and Erbitux (cetuximab). In the lot-to-lot glycan structure variation test for Herceptin or Avastin, around 30 glycan structures on Fc region of antibody drugs were relatively quantified in 30-min analysis, revealing that at most 10% increase or decrease of several glycoforms were observed in both drugs (n = 4 for each of four lots). In the case of Avastin glycan profiling, glycans on both Fab and Fc

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

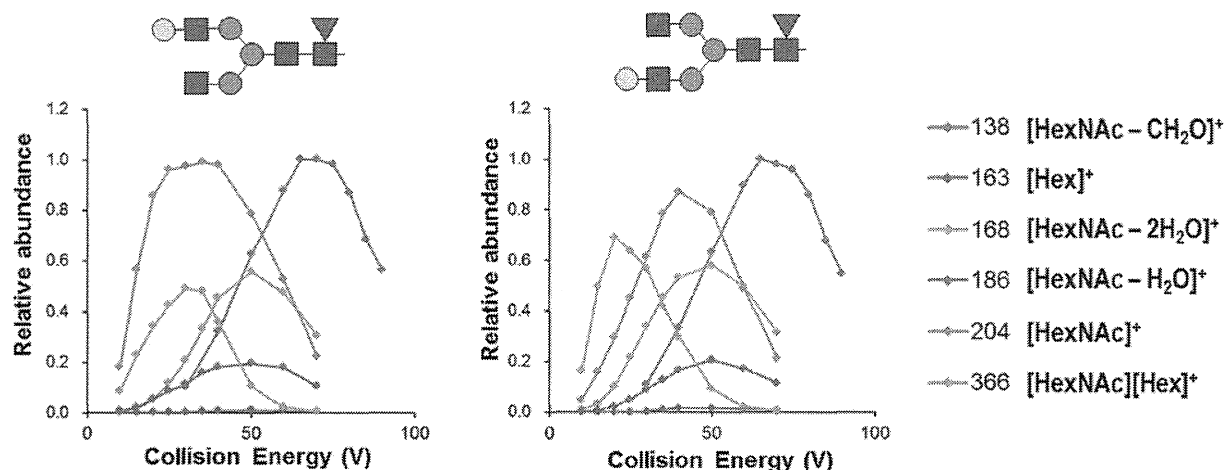


Figure 4. Quantitative evaluation of site-specific glycan structure microheterogeneities by Erexim technology. Representative Erexim curves for a pair of structural isomers detected on IgG molecule are shown. Six oxonium ions listed on the right side were quantitatively monitored by MRM mode. The patterns of collision energy-dependent ion yield curves for $m/z = 204$ and 366 were clearly distinguishable between α1,6-linked galactose (left) and α1,3-linked galactose (right) at the nonreducing termini.

regions were simultaneously quantified in a single Erexim analysis. The result showed that most of glycans on Fab region were nonhuman-type glycan structures possessing *N*-glycolylneuraminic acids (Neu5Gc) or Gal (α1–3) Gal structures, whereas Fc region had conserved nonimmunogenic glycan structures. The LOD and dynamic range of this technology were 30 attomole and more than four orders, respectively. Since the required sample preparation prior to mass spectrometric analysis was only usual trypsin digestion, the Erexim procedure would be appropriate for automated high-throughput analysis. Thus, this technology has enough potential to be applied to routine evaluation of drug quality, safety, and potency, for which extremely high reproducibility, quantitative capability, and throughput. Meanwhile, our Erexim profiling technique is promising technology allowing rapid and site-specific validation of extracted glycan structure-targeted tumor marker candidates using multiple crude specimens.

4 Toward industrialization and approval of carbohydrate-targeting biomarkers

Biomarkers developed for commercial use and regulatory approval must be required to present data supporting validity and clinical utility. According to the FDA Guidance for industry: Pharmacogenomic data submissions 2005, a valid biomarker is a biomarker that is measured in an analytical test system with "well-established performance characteristics" and for which there is an "established scientific framework or body of evidence that elucidates the clinical significance of test results." In case of carbohydrate-targeting biomarkers for cancer early detection or prognosis, the developed devices need to be approved as in vitro diagnostics.

Toward that purpose, a couple of key requirements have to be fulfilled. First, development of high-throughput, easy-to-use, reproducible, and not so expensive diagnostic devices would be required, which are suitable for widespread clinical use. As described in Section 3.1, lectin-antibody sandwich ELISA system could be the first choice satisfying all of these criteria. However, in some cases it would be difficult to achieve sensitive detection for low abundant serum glycoproteins because of insufficient specificity and affinity of lectins for the glycan epitopes. The Erexim technology (Section 3.2) can provide extremely high sensitivity, specificity, and reproducibility for the quantitative assessment of site-specific glycan structure disorders, whereas the high-spec mass spectrometer and proficient skills of LC/MS/MS operation are required.

The other requirement toward the approval of diagnostics is biological evidence explaining how and why the targeted glycan structure alterations would occur on a particular serum glycoprotein in cancer patients. It is often complicated to figure out the origin of biomarker glycoproteins and biochemical mechanisms of abnormal oligosaccharide generation, especially when investigating for glycoproteins identified from serum shotgun proteomics-driven biomarker screening. To present scientifically strong evidences for carbohydrate-targeting biomarker candidates, further basic and detailed glycobiological studies will be necessary, including analysis of cellular glycan biosynthesis pathways, comprehensive expression analysis of glycosyltransferases, or confirmation of glycoform changes in cancer tissues. Future development of more sensitive, high-throughput, and site-specific glycan structure profiling technologies are also fundamental to facilitate clinical application of numerous carbohydrate-targeting biomarker candidates reported previously on papers.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

The authors have declared no conflict of interest.

5 References

[1] Anderson, N. L., The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clin. Chem.* 2010, *56*, 177–185.

[2] Anderson, N. L., Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* 2002, *1*, 845–867.

[3] Ueda, K., Saichi, N., Takami, S., Kang, D. et al., A comprehensive peptidome profiling technology for the identification of early detection biomarkers for lung adenocarcinoma. *PLoS One* 2011, *6*, e18567.

[4] Schulz-Knappe, P., Zucht, H. D., Heine, G., Jurgens, M. et al., Peptidomics: the comprehensive analysis of peptides in complex biological mixtures. *Comb. Chem. High Throughput Screen.* 2001, *4*, 207–217.

[5] Machtejevas, E., Andrecht, S., Lubda, D., Unger, K. K., Monolithic silica columns of various format in automated sample clean-up/multidimensional liquid chromatography/mass spectrometry for peptidomics. *J. Chromatogr. A* 2007, *1144*, 97–101.

[6] Fogli, A., Bulet, P., Peptidomics analysis of lymphoblastoid cell lines. *Methods Mol. Biol.* 2010, *615*, 247–257.

[7] Overall, C. M., Dean, R. A., Degradomics: systems biology of the protease web. Pleiotropic roles of MMPs in cancer. *Cancer Metastasis Rev.* 2006, *25*, 69–75.

[8] Doucet, A., Butler, G. S., Rodriguez, D., Prudova, A. et al., Metadegradomics: toward in vivo quantitative degradomics of proteolytic post-translational modifications of the cancer proteome. *Mol. Cell. Proteomics* 2008, *7*, 1925–1951.

[9] Butler, G. S., Dean, R. A., Smith, D., Overall, C. M., Membrane protease degradomics: proteomic identification and quantification of cell surface protease substrates. *Methods Mol. Biol.* 2009, *528*, 159–176.

[10] Huesgen, P. F., Overall, C. M., N- and C-terminal degradomics: new approaches to reveal biological roles for plant proteases from substrate identification. *Physiol. Plant.* 2012, *145*, 5–17.

[11] Wejda, M., Impens, F., Takahashi, N., Van Damme, P. et al., Degradomics reveals that cleavage specificity profiles of caspase-2 and effector caspases are alike. *J. Biol. Chem.* 2012, *287*, 33983–33995.

[12] Wehr, A. Y., Hwang, W. T., Blair, I. A., Yu, K. H., Relative quantification of serum proteins from pancreatic ductal adenocarcinoma patients by stable isotope dilution liquid chromatography-mass spectrometry. *J. Proteome Res.* 2012, *11*, 1749–1758.

[13] Shi, T., Zhou, J. Y., Gritsenko, M. A., Hossain, M. et al., IgY14 and SuperMix immunoaffinity separations coupled with liquid chromatography-mass spectrometry for human plasma proteomics biomarker discovery. *Methods* 2012, *56*, 246–253.

[14] Koutroukides, T. A., Guest, P. C., Leweke, F. M., Bailey, D. M. et al., Characterization of the human serum depletome by label-free shotgun proteomics. *J. Sep. Sci.* 2011, *34*, 1621–1626.

[15] Beer, L. A., Tang, H. Y., Sriswasdi, S., Barnhart, K. T. et al., Systematic discovery of ectopic pregnancy serum biomarkers using 3-D protein profiling coupled with label-free quantitation. *J. Proteome Res.* 2011, *10*, 1126–1138.

[16] Berven, F. S., Ahmad, R., Clauser, K. R., Carr, S. A., Optimizing performance of glycopeptide capture for plasma proteomics. *J. Proteome Res.* 2010, *9*, 1706–1715.

[17] Jagtap, P., Bandhakavi, S., Higgins, L., McGowan, T. et al., Workflow for analysis of high mass accuracy salivary data set using MaxQuant and ProteinPilot search algorithm. *Proteomics* 2012, *12*, 1726–1730.

[18] Milan, E., Lazzari, C., Anand, S., Floriani, I. et al., SAA1 is over-expressed in plasma of non small cell lung cancer patients with poor outcome after treatment with epidermal growth factor receptor tyrosine-kinase inhibitors. *J. Proteomics* 2012, *76*, 91–101.

[19] Righetti, P. G., Boschetti, E., The ProteoMiner and the FortyNiners: searching for gold nuggets in the proteomic arena. *Mass Spectrom. Rev.* 2008, *27*, 596–608.

[20] Boschetti, E., Righetti, P. G., The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. *J. Proteomics* 2008, *71*, 255–264.

[21] Hartwig, S., Czibere, A., Kotzka, J., Passlack, W. et al., Combinatorial hexapeptide ligand libraries (ProteoMiner): an innovative fractionation tool for differential quantitative clinical proteomics. *Arch. Physiol. Biochem.* 2009, *115*, 155–160.

[22] Fonslow, B. R., Carvalho, P. C., Academia, K., Freeby, S. et al., Improvements in proteomic metrics of low abundance proteins through proteome equalization using ProteoMiner prior to MudPIT. *J. Proteome Res.* 2011, *10*, 3690–3700.

[23] Hekmat, O., He, S., Warren, R. A., Withers, S. G., A mechanism-based ICAT strategy for comparing relative expression and activity levels of glycosidases in biological systems. *J. Proteome Res.* 2008, *7*, 3282–3292.

[24] Kang, X., Sun, L., Guo, K., Shu, H. et al., Serum protein biomarkers screening in HCC patients with liver cirrhosis by ICAT-LC-MS/MS. *J. Cancer Res. Clin. Oncol.* 2010, *136*, 1151–1159.

[25] Ramus, C., Gonzalez de Peredo, A., Dahout, C., Gallagher, M. et al., An optimized strategy for ICAT quantification of membrane proteins. *Mol. Cell. Proteomics* 2006, *5*, 68–78.

[26] Sethuraman, M., McComb, M. E., Huang, H., Huang, S. et al., Isotope-coded affinity tag (ICAT) approach to redox proteomics: identification and quantitation of oxidant-sensitive cysteine thiols in complex protein mixtures. *J. Proteome Res.* 2004, *3*, 1228–1233.

[27] Yan, W., Lee, H., Deutsch, E. W., Lazaro, C. A. et al., A dataset of human liver proteins identified by protein profiling via isotope-coded affinity tag (ICAT) and tandem mass spectrometry. *Mol. Cell. Proteomics* 2004, *3*, 1039–1041.

[28] Li, J., Steen, H., Gygi, S. P., Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. *Mol. Cell. Proteomics* 2003, *2*, 1198–1204.

[29] Ueda, K., Katagiri, T., Shimada, T., Irie, S. et al., Comparative profiling of serum glycoproteome by sequential

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

purification of glycoproteins and 2-nitrobenzensulfenyl (NBS) stable isotope labeling: a new approach for the novel biomarker discovery for cancer. *J. Proteome Res.* 2007, *6*, 3475–3483.

[30] Iida, T., Kuyama, H., Watanabe, M., Toda, C. et al., Rapid and efficient MALDI-TOF MS peak detection of 2-nitrobenzenesulfenyl-labeled peptides using the combination of HPLC and an automatic spotting apparatus. *J. Biomol. Tech.* 2006, *17*, 333–341.

[31] Ou, K., Kesuma, D., Ganesan, K., Yu, K. et al., Quantitative profiling of drug-associated proteomic alterations by combined 2-nitrobenzenesulfenyl chloride (NBS) isotope labeling and 2DE/MS identification. *J. Proteome Res.* 2006, *5*, 2194–2206.

[32] Matsuo, E., Toda, C., Watanabe, M., Ojima, N. et al., Selective detection of 2-nitrobenzenesulfenyl-labeled peptides by matrix-assisted laser desorption/ionization-time of flight mass spectrometry using a novel matrix. *Proteomics* 2006, *6*, 2042–2049.

[33] Matsuo, E., Toda, C., Watanabe, M., Iida, T. et al., Improved 2-nitrobenzenesulfenyl method: optimization of the protocol and improved enrichment for labeled peptides. *Rapid Commun. Mass Spectrom.* 2006, *20*, 31–38.

[34] Kuyama, H., Watanabe, M., Toda, C., Ando, E. et al., An approach to quantitative proteome analysis by labeling tryptophan residues. *Rapid Commun. Mass Spectrom.* 2003, *17*, 1642–1650.

[35] Kurogochi, M., Amano, M., Fumoto, M., Takimoto, A. et al., Reverse glycoblotting allows rapid-enrichment glycoproteomics of biopharmaceuticals and disease-related biomarkers. *Angew. Chem.Int. Ed. Engl.* 2007, *46*, 8808–8813.

[36] Kurogochi, M., Matsushista, T., Amano, M., Furukawa, J. et al., Sialic acid-focused quantitative mouse serum glycoproteomics by multiple reaction monitoring assay. *Mol. Cell. Proteomics* 2010, *9*, 2354–2368.

[37] Ziegler, A., Cerciello, F., Bigosch, C., Bausch-Fluck, D. et al., Proteomic surfaceome analysis of mesothelioma. *Lung Cancer* 2012, *75*, 189–196.

[38] Gundry, R. L., Riordon, D. R., Tarasova, Y., Chuppa, S. et al., A cell surfaceome map for immunophenotyping and sorting pluripotent stem cells. *Mol. Cell. Proteomics* 2012, *11*, 303–316.

[39] Danzer, C., Eckhardt, K., Schmidt, A., Fankhauser, N. et al., Comprehensive description of the N-glycoproteome of mouse pancreatic beta-cells and human islets. *J. Proteome Res.* 2012, *11*, 1598–1608.

[40] Bausch-Fluck, D., Hofmann, A., Wollscheid, B., Cell surface capturing technologies for the surfaceome discovery of hepatocytes. *Methods Mol. Biol.* 2012, *909*, 1–16.

[41] Kovarova, H., Gadher, S. J., Wollscheid, B., Focus on stem cell proteomics. *Proteomics* 2011, *11*, 3943–3945.

[42] Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R. et al., Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol.* 2009, *27*, 378–386.

[43] Liu, L., Zhang, Y., Zhang, L., Yan, G. et al., Highly specific revelation of rat serum glycopeptidome by boronic acid-functionalized mesoporous silica. *Anal. Chim. Acta* 2012, *753*, 64–72.

[44] Xu, Y., Zhang, L., Lu, H., Yang, P., On-plate enrichment of glycopeptides by using boronic acid functionalized gold-coated Si wafer. *Proteomics* 2010, *10*, 1079–1086.

[45] Tang, J., Liu, Y., Qi, D., Yao, G. et al., On-plate-selective enrichment of glycopeptides using boronic acid-modified gold nanoparticles for direct MALDI-QIT-TOF MS analysis. *Proteomics* 2009, *9*, 5046–5055.

[46] Zhu, J., Wang, F., Chen, R., Cheng, K. et al., Centrifugation assisted microreactor enables facile integration of trypsin digestion, hydrophilic interaction chromatography enrichment, and on-column deglycosylation for rapid and sensitive N-glycoproteome analysis. *Anal. Chem.* 2012, *84*, 5146–5153.

[47] Xiong, Z., Zhao, L., Wang, F., Zhu, J. et al., Synthesis of branched PEG brushes hybrid hydrophilic magnetic nanoparticles for the selective enrichment of N-linked glycopeptides. *Chem. Commun.* 2012, *48*, 8138–8140.

[48] Kuo, C. W., Wu, I. L., Hsiao, H. H., Khoo, K. H., Rapid glycopeptide enrichment and N-glycosylation site mapping strategies based on amine-functionalized magnetic nanoparticles. *Anal. Bioanal. Chem.* 2012, *402*, 2765–2776.

[49] Dam, T. K., Gerken, T. A., Brewer, C. F., Thermodynamics of multivalent carbohydrate-lectin cross-linking interactions: importance of entropy in the bind and jump mechanism. *Biochemistry* 2009, *48*, 3822–3827.

[50] Ueda, K., Takami, S., Saichi, N., Daigo, Y. et al., Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Mol. Cell. Proteomics* 2010, *9*, 1819–1828.

[51] Kaji, H., Saito, H., Yamauchi, Y., Shinkawa, T. et al., Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat. Biotechnol.* 2003, *21*, 667–672.

[52] Kaji, H., Shikanai, T., Sasaki-Sawa, A., Wen, H. et al., Large-scale identification of N-glycosylated proteins of mouse tissues and construction of a glycoprotein database, GlycoProtDB. *J. Proteome Res.* 2012, *11*, 4553–4566.

[53] Liu, T. W., Kaji, H., Togayachi, A., Ito, H. et al., A chemoenzymatic approach toward the identification of fucosylated glycoproteins and mapping of N-glycan sites. *Glycobiology* 2012, *22*, 630–637.

[54] Narimatsu, H., Sawaki, H., Kuno, A., Kaji, H. et al., A strategy for discovery of cancer glyco-biomarkers in serum using newly developed technologies for glycoproteomics. *FEBS J.* 2010, *277*, 95–105.

[55] Sugahara, D., Kaji, H., Sugihara, K., Asano, M. et al., Large-scale identification of target proteins of a glycosyltransferase isozyme by lectin-IGOT-LC/MS, an LC/MS-based glycoproteomic approach. *Sci. Rep.* 2012, *2*, 680.

[56] Han, H., Stapels, M., Ying, W., Yu, Y. et al., Comprehensive characterization of the N-glycosylation status of CD44s by use of multiple mass spectrometry-based techniques. *Anal. Bioanal. Chem.* 2012, *404*, 373–388.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[57] Zhang, S., Liu, X., Kang, X., Sun, C. et al., iTRAQ plus 18O: a new technique for target glycoprotein analysis. *Talanta* 2012, *91*, 122–127.

[58] Kubota, K., Sato, Y., Suzuki, Y., Goto-Inoue, N. et al., Analysis of glycopeptides using lectin affinity chromatography with MALDI-TOF mass spectrometry. *Anal. Chem.* 2008, *80*, 3693–3698.

[59] Snovida, S. I., Bodnar, E. D., Viner, R., Saba, J. et al., A simple cellulose column procedure for selective enrichment of glycopeptides and characterization by nano LC coupled with electron-transfer and high-energy collisional-dissociation tandem mass spectrometry. *Carbohydr. Res.* 2010, *345*, 792–801.

[60] Wada, Y., Tajiri, M., Yoshida, S., Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* 2004, *76*, 6560–6565.

[61] Davies, M., Smith, K. D., Harbin, A. M., Hounsell, E. F., High-performance liquid chromatography of oligosaccharide alditols and glycopeptides on a graphitized carbon column. *J. Chromatogr.* 1992, *609*, 125–131.

[62] Davies, M. J., Smith, K. D., Carruthers, R. A., Chai, W. et al., Use of a porous graphitised carbon column for the high-performance liquid chromatography of oligosaccharides, alditols and glycopeptides with subsequent mass spectrometry analysis. *J. Chromatogr.* 1993, *646*, 317–326.

[63] Fan, J. Q., Kondo, A., Kato, I., Lee, Y. C., High-performance liquid chromatography of glycopeptides and oligosaccharides on graphitized carbon columns. *Anal. Biochem.* 1994, *219*, 224–229.

[64] Thaysen-Andersen, M., Mysling, S., Hojrup, P., Site-specific glycoprofiling of N-linked glycopeptides using MALDI-TOF MS: strong correlation between signal strength and glycoform quantities. *Anal. Chem.* 2009, *81*, 3933–3943.

[65] Lam, M. P., Lau, E., Siu, S. O., Ng, D. C. et al., Online combination of reversed-phase/reversed-phase and porous graphitic carbon liquid chromatography for multicomponent separation of proteomics and glycoproteomics samples. *Electrophoresis* 2011, *32*, 2930–2940.

[66] Selman, M. H., McDonnell, L. A., Palmblad, M., Ruhaak, L. R. et al., Immunoglobulin G glycopeptide profiling by matrix-assisted laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 2010, *82*, 1073–1081.

[67] Zeng, Z., Hincapie, M., Pitteri, S. J., Hanash, S. et al., A proteomics platform combining depletion, multi-lectin affinity chromatography (M-LAC), and isoelectric focusing to study the breast cancer proteome. *Anal. Chem.* 2011, *83*, 4845–4854.

[68] Yang, Z., Hancock, W. S., Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. *J. Chromatogr. A* 2004, *1053*, 79–88.

[69] Yang, Z., Hancock, W. S., Monitoring glycosylation pattern changes of glycoproteins using multi-lectin affinity chromatography. *J. Chromatogr. A* 2005, *1070*, 57–64.

[70] Dayarathna, M. K., Hancock, W. S., Hincapie, M., A two step fractionation approach for plasma proteomics using immunodepletion of abundant proteins and multi-lectin affinity chromatography: application to the analysis of obesity, diabetes, and hypertension diseases. *J. Sep. Sci.* 2008, *31*, 1156–1166.

[71] Kullolli, M., Hancock, W. S., Hincapie, M., Preparation of a high-performance multi-lectin affinity chromatography (HP-M-LAC) adsorbent for the analysis of human plasma glycoproteins. *J. Sep. Sci.* 2008, *31*, 2733–2739.

[72] Zheng, X., Wu, S. L., Hincapie, M., Hancock, W. S., Study of the human plasma proteome of rheumatoid arthritis. *J. Chromatogr. A* 2009, *1216*, 3538–3545.

[73] Qiu, R., Regnier, F. E., Use of multidimensional lectin affinity chromatography in differential glycoproteomics. *Anal. Chem.* 2005, *77*, 2802–2809.

[74] Hang, H. C., Yu, C., Kato, D. L., Bertozzi, C. R., A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 14846–14851.

[75] Laughlin, S. T., Bertozzi, C. R., Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. *Nat. Protoc.* 2007, *2*, 2930–2944.

[76] Slade, P. G., Hajivandi, M., Bartel, C. M., Gorfien, S. F., Identifying the CHO secretome using mucin-type O-linked glycosylation and Click-chemistry. *J. Proteome Res.* 2012, *11*, 6175–6186.

[77] Dube, D. H., Prescher, J. A., Quang, C. N., Bertozzi, C. R., Probing mucin-type O-linked glycosylation in living animals. *Proc. Natl. Acad. Sci. USA* 2006, *103*, 4819–4824.

[78] Laughlin, S. T., Baskin, J. M., Amacher, S. L., Bertozzi, C. R., In vivo imaging of membrane-associated glycans in developing zebrafish. *Science* 2008, *320*, 664–667.

[79] Hart, C., Chase, L. G., Hajivandi, M., Agnew, B., Metabolic labeling and click chemistry detection of glycoprotein markers of mesenchymal stem cell differentiation. *Methods Mol. Biol.* 2011, *698*, 459–484.

[80] Foo, J. Y., Wan, Y., Kostner, K., Arivalagan, A. et al., NT-ProBNP levels in saliva and its clinical relevance to heart failure. *PLoS One* 2012, *7*, e48452.

[81] Pfluger, M., Kapuscik, A., Lucas, R., Koppensteiner, A. et al., A combined impedance and AlphaLISA-based approach to identify anti-inflammatory and barrier-protective compounds in human endothelium. *J. Biomol. Screen.* 2012, *18*, 67–74.

[82] Peters, C. D., Jespersen, B., Norregaard, R., AlphaLISA versus ELISA-based detection of interleukin 18 in healthy subjects and patients with end-stage renal disease. *Scand. J. Clin. Lab. Invest.* 2012, *72*, 583–592.

[83] Cosentino, G., AlphaLISA assays to improve the vaccine development process. *Dev. Biol.* 2012, *134*, 107–111.

[84] Waller, H., Chatterji, U., Gallay, P., Parkinson, T. et al., The use of AlphaLISA technology to detect interaction between hepatitis C virus-encoded NS5A and cyclophilin A. *J. Virol. Methods* 2010, *165*, 202–210.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[85] Biancotto, A., Feng, X., Langweiler, M., Young, N. S. et al., Effect of anticoagulants on multiplexed measurement of cytokine/chemokines in healthy subjects. *Cytokine* 2012, *60*, 438–446.

[86] Kim, Y. W., Bae, S. M., Lim, H., Kim, Y. J. et al., Development of multiplexed bead-based immunoassays for the detection of early stage ovarian cancer using a combination of serum biomarkers. *PLoS One* 2012, *7*, e44960.

[87] Rinewalt, D., Shersher, D. D., Daly, S., Fhied, C. et al., Development of a serum biomarker panel predicting recurrence in stage I non-small cell lung cancer patients. *J. Thorac. Cardiovasc. Surg.* 2012, *144*, 1344–1351.

[88] Tait, B. D., Hudson, F., Cantwell, L., Brewin, G. et al., Review article: Luminex technology for HLA antibody detection in organ transplantation. *Nephrology* 2009, *14*, 247–254.

[89] Picascia, A., Infante, T., Napoli, C., Luminex and antibody detection in kidney transplantation. *Clin. Exp. Nephrol.* 2012, *16*, 373–381.

[90] Drouin, J., Izaguirre, C. A., Patenaude, P., Quantitation of cell membrane glycoproteins in pathological conditions using a lectin-bound enzyme-linked immunosorbent assay (ELISA). Application to human platelets in the Bernard-Soulier syndrome. *J. Immunol. Methods* 1988, *110*, 217–223.

[91] Tojo, M., Shibata, N., Osanai, T., Mikami, T. et al., Sandwich enzyme-linked immunosorbent assay of D-mannans of *Candida albicans* NIH A-207 and NIH B-792 strains using concanavalin A and polyclonal rabbit anti-*C. albicans* antisera. *Carbohydr. Res.* 1991, *213*, 325–330.

[92] Parker, N., Lectin/Antibody "sandwich" ELISA for quantification of circulating mucin as a diagnostic test for pancreatic cancer. *Methods Mol. Med.* 1998, *9*, 249–253.

[93] Parker, N., Makin, C. A., Ching, C. K., Eccleston, D. et al., A new enzyme-linked lectin/mucin antibody sandwich assay (CAM 17.1/WGA) assessed in combination with CA 19-9 and peanut lectin binding assay for the diagnosis of pancreatic cancer. *Cancer* 1992, *70*, 1062–1068.

[94] Yiannakou, J. Y., Newland, P., Calder, F., Kingsnorth, A. N. et al., Prospective study of CAM 17.1/WGA mucin assay for serological diagnosis of pancreatic cancer. *Lancet* 1997, *349*, 389–392.

[95] Takeda, Y., Shinzaki, S., Okudo, K., Moriwaki, K. et al., Fucosylated haptoglobin is a novel type of cancer biomarker linked to the prognosis after an operation in colorectal cancer. *Cancer* 2012, *118*, 3036–3043.

[96] Okuyama, N., Ide, Y., Nakano, M., Nakagawa, T. et al., Fucosylated haptoglobin is a novel marker for pancreatic cancer: a detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation. *Int. J. Cancer* 2006, *118*, 2803–2808.

[97] Miyoshi, E., Shinzaki, S., Moriwaki, K., Matsumoto, H., Identification of fucosylated haptoglobin as a novel tumor marker for pancreatic cancer and its possible application for a clinical diagnostic test. *Methods Enzymol.* 2010, *478*, 153–164.

[98] Matsumoto, H., Shinzaki, S., Narisada, M., Kawamoto, S. et al., Clinical application of a lectin-antibody ELISA to measure fucosylated haptoglobin in sera of patients with pancreatic cancer. *Clin. Chem. Lab. Med.* 2010, *48*, 505–512.

[99] Toyama, A., Nakagawa, H., Matsuda, K., Sato, T. A. et al., Quantitative structural characterization of local N-glycan microheterogeneity in therapeutic antibodies by energy-resolved oxonium ion monitoring. *Anal. Chem.* 2012, *84*, 9655–9662.

[100] Medzihradszky, K. F., Characterization of site-specific N-glycosylation. *Methods Mol. Biol.* 2008, *446*, 293–316.

[101] Scott, N. E., Parker, B. L., Connolly, A. M., Paulech, J. et al., Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. *Mol. Cell. Proteomics* 2011, *10*, M000031-MCP000201.

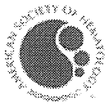
[102] Lemoine, J., Fortin, T., Salvador, A., Jaffuel, A. et al., The current status of clinical proteomics and the use of MRM and MRM(3) for biomarker validation. *Expert Rev. Mol. Diagn.* 2012, *12*, 333–342.

[103] Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L. et al., Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J. Chromatogr. B* 2009, *877*, 1229–1239.

[104] Shi, T., Su, D., Liu, T., Tang, K. et al., Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. *Proteomics* 2012, *12*, 1074–1092.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.



blood

2013 121: 4340-4347
doi:10.1182/blood-2012-08-446922 originally published online March 28, 2013

Preapoptotic protease calpain-2 is frequently suppressed in adult T-cell leukemia

Makoto Ishihara, Natsumi Araya, Tomoo Sato, Ayako Tatsuguchi, Naomi Saichi, Atae Utsunomiya, Yusuke Nakamura, Hidewaki Nakagawa, Yoshihisa Yamano and Koji Ueda

Updated information and services can be found at:
http://bloodjournal.hematologylibrary.org/content/121/21/4340.full.html

Articles on similar topics can be found in the following Blood collections
Lymphoid Neoplasia (1685 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.
Copyright 2011 by The American Society of Hematology; all rights reserved.

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

LYMPHOID NEOPLASIA

Preapoptotic protease calpain-2 is frequently suppressed in adult T-cell leukemia

Makoto Ishihara,[1] Natsumi Araya,[2] Tomoo Sato,[2] Ayako Tatsuguchi,[1] Naomi Saichi,[1] Atae Utsunomiya,[3] Yusuke Nakamura,[4] Hidewaki Nakagawa,[1] Yoshihisa Yamano,[2] and Koji Ueda[1]

[1]Laboratory for Biomarker Development, Center of Genomic Medicine, RIKEN, Tokyo, Japan; [2]Department of Molecular Medical Science, Institute of Medical Science, St. Marianna University School of Medicine, Kawasaki, Japan; [3]Department of Hematology, Imamura Bun-in Hospital, Kagoshima, Japan; and [4]Section of Hematology/Oncology, Department of Medicine Faculty, The University of Chicago, Chicago, IL

Key Points

- Proteome-wide analysis of HTLV-1–infected T cells identified 17 biomarker proteins for the diagnosis of ATL or HAM/TSP patients.

Adult T-cell leukemia (ATL) is one of the most aggressive hematologic malignancies caused by human T-lymphotropic virus type 1 (HTLV-1) infection. The prognosis of ATL is extremely poor; however, effective strategies for diagnosis and treatment have not been established. To identify novel therapeutic targets and diagnostic markers for ATL, we employed focused proteomic profiling of the $CD4^+CD25^+CCR4^+$ T-cell subpopulation in which HTLV-1–infected cells were enriched. Comprehensive quantification of 14 064 peptides and subsequent 2-step statistical analysis using 29 cases (6 uninfected controls, 5 asymptomatic carriers, 9 HTLV-1–associated myelopathy/tropical spastic paraparesis patients, 9 ATL patients) identified 91 peptide determinants that statistically classified 4 clinical groups with an accuracy rate of 92.2% by cross-validation test. Among the identified 17 classifier proteins, α-II spectrin was drastically accumulated in infected T cells derived from ATL patients, whereas its digestive protease calpain-2 (CAN2) was significantly downregulated. Further cell cycle analysis and cell growth assay revealed that rescue of CAN2 activity by overexpressing constitutively active CAN2 (Δ_{19}CAN2) could induce remarkable cell death on ATL cells accompanied by reduction of α-II spectrin. These results support that proteomic profiling of HTLV-1–infected T cells could provide potential diagnostic biomarkers and an attractive resource of therapeutic targets for ATL. (*Blood*. 2013;121(21):4340-4347)

Introduction

Human T-lymphotropic virus type 1 (HTLV-1) is a human retrovirus that is the pathogenic agent of HTLV-1–associated diseases, such as adult T-cell leukemia (ATL) and HTLV-1–associated myelopathy/tropical spastic paraparesis (HAM/TSP). Recent epidemiological studies revealed that HTLV-1 is endemic mainly in Japan, the Caribbean basin, Iran, Africa, South America, and the Melanesian islands.[1] Other estimates have shown that 20 million to 30 million people worldwide are infected with HTLV-1.[2] The infection is followed by a prolonged asymptomatic phase of 20 to 30 years, and 2% to 5% of the infected individuals develop ATL during their lifetime.[3] ATL is one of the most aggressive hematologic malignancies characterized by increased numbers of lymphocytes with multilobulated nuclei, so-called flower cells, in blood circulation. The prognosis is severe with the median overall survival period and 5-year survival rate of ATL patients of 7 months and 20%, respectively.[4] Recently, humanized anti-CCR4 (KW-0761) therapeutic antibody achieved a great improvement in ATL treatment in a phase 3 study. However, the disease control rate was restricted to 50%, and long-term prognosis has yet to be known.[5] For future improvements in the management of ATL, novel biomarkers for early diagnosis are urgently needed for early therapeutic intervention.

To date, comprehensive genomic or proteomic studies using $CD4^+$ T cells have been performed for this purpose,[6-9] but reproducibility and reliability of quantification results in the discovery phase were uncertain due to the diverse individual variety of HTLV-1–infected cell contents in $CD4^+$ T cells. To overcome the etiologic variety of samples, we focused on the $CD4^+CD25^+CCR4^+$ T-cell subpopulation since Yamano et al[10] recently revealed that HTLV-1 preferentially infected $CD4^+CD25^+CCR4^+$ T cells in both ATL and HAM/TSP patients. By targeting $CD4^+CD25^+CCR4^+$ T cells, we here provide the first quantitative proteome map illustrating molecular disorders in pathogenic human T cells directly associated with the onset or progression of ATL. The comprehensive and comparative interpretation of total proteome in infected cells, especially between asymptomatic HTLV-1 carriers and ATL patients, could immediately lead to specific candidates for biomarkers and drugs.

Another challenge to emphasize in this study is our recently established proteomic profiling technologies. It is indisputable that the greater the number of clinical samples analyzed, the more confidently statistical analysis can be undertaken in order to identify diagnostic markers and druggable targets. Despite this fact, previous proteomics reports could not provide high-throughput quantitative methodologies that were sufficient for dealing with even more than 10 clinical samples, excepting a study utilizing a surface enhanced laser desorption/ionization time of flight mass spectrometer. Although the surface enhanced laser desorption/ionization time of flight method drastically improved the performance in both quantification and throughput, allowing relative quantification

Submitted August 1, 2012; accepted March 25, 2013. Prepublished online as *Blood* First Edition paper, March 28, 2013; DOI 10.1182/blood-2012-08-446922.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2013 by The American Society of Hematology