Table 1: Clinicopathological characteristics of individuals examined in this study.

		All cases ($n = 162$)						Cases analyzed by 2DICAL ($n = 40$)		
		PCa* ($n = 54$)	BPH ($n = 22$)	Prostatitis ($n = 6$)	RCC ($n = 20$)	Healthy ($n = 60$)	P value	PCa* ($n = 25$)	Healthy ($n = 15$)	P value
Age	(mean ± SD)	66. 3 ± 7.3	72.3 ± 7.5	65.8 ± 5.9	63.8 ± 8.5	67.2 ± 8.8	0.544****	67.0 ± 7.2	65.1 ± 8.0	0.440****
Stage	I	0						0		
	II	50						22		
	III	2						2		
	IV	2						1		
TNM classification**	T1cN0M0	25						10		
	T2aN0M0	15						7		
	T2bN0M0	2						1		
	T2cN0M0	8						4		
	T3aN0M0	2						2		
	T4N0M1b	1						1		
	T4N1M1b	1						0		
PSA***	ng/mL									
	<4	2	11	4		51		2	12	
	4≦, ≦10	30	10	2		9	<0.001*****	14	3	<0.001*****
	10<	22	1	0		0		9	0	
Gleason score	5	2						1		
	6	15						7		
	7	25						12		
	8	8						4		
	9	2						1		
	10	2						0		

*Abbreviation: PCa: prostate cancer; **according to TNM Classification of Malignant Tumors (International Union Against Cancer), 6th Edition; ***measured by Tandem R kit before the first treatment; ****Welch's t-test; *****Fisher's exact test.

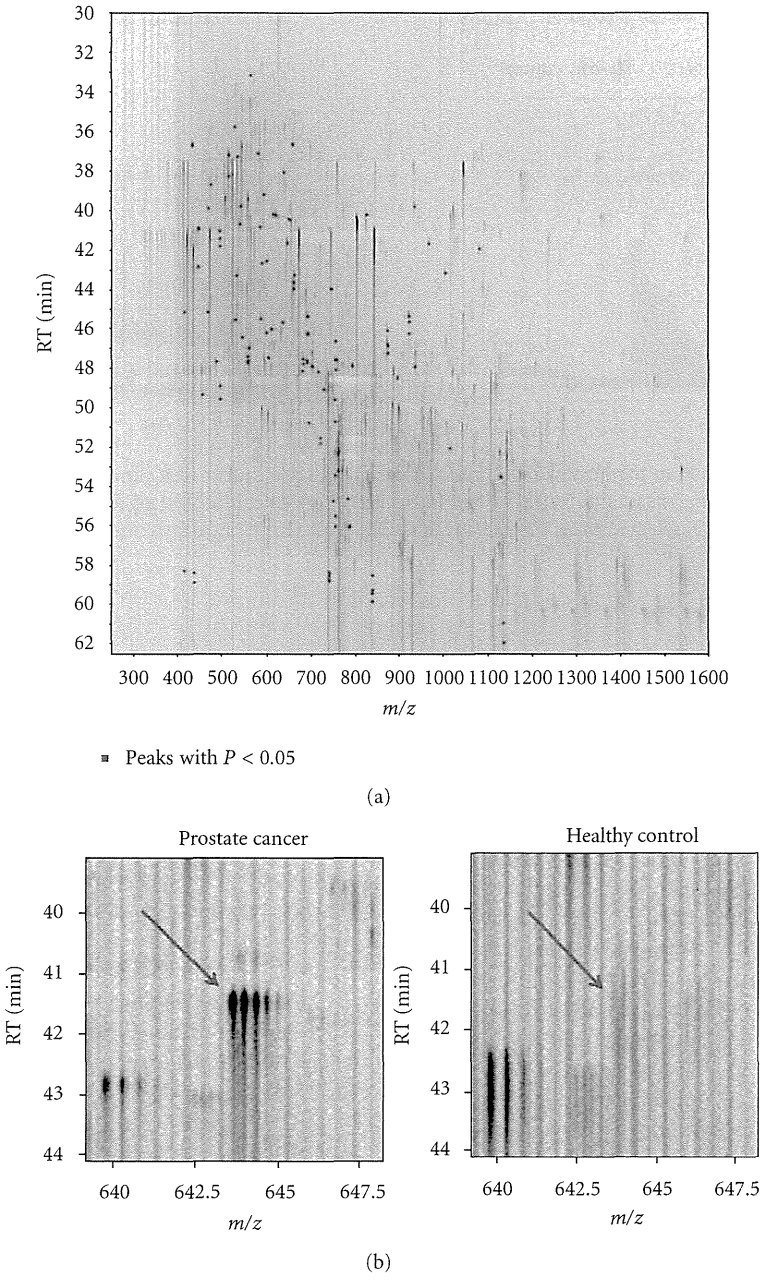

FIGURE 1: (a) Two-dimensional display of all MS peaks. The 117 MS peaks whose mean intensities significantly differed between prostate cancer patients and healthy controls ($P < 0.05$, Welch's t-test) are highlighted in red. RT: retention time. (b) 2DICAL images of peak ID 396 in a representative prostate cancer patient (left) and a healthy control (right).

OPD solution for 10 min, absorbance was measured at 490 nm using EnSpire AlphaPLUS (PerkinElmer, Waltham, MA).

2.7. Cell Lines. The human prostate cancer cell line 22Rv1 was purchased from Riken BRC Cell Bank (Tsukuba, Japan) and cultured in Roswell Park Memorial Institute Medium 1640 supplemented with 10% fetal bovine serum. Normal human prostate epithelial cells (CC-3165 PrEBM) were purchased from Lonza (Basel, Switzerland). The cells were cultured at 37°C under 5% CO_2. The culture medium was changed every 3 days. After incubation of cell lines for 48 h, the cultured conditioned media were collected, filtered through a sterile 0.22-mm filter unit, and concentrated 100-fold by freeze drying. A 10 μL sample of concentrated medium was used for Western blot analysis.

2.8. Immunohistochemistry. Cells were seeded on BioCoat collagen I-coated 8-well culture slides (Becton Dickinson Labware, Bedford, MA) and cultured overnight. The cells were washed twice with PBS (pH 7.2) and fixed with 4% paraformaldehyde in PBS for 20 min at room temperature.

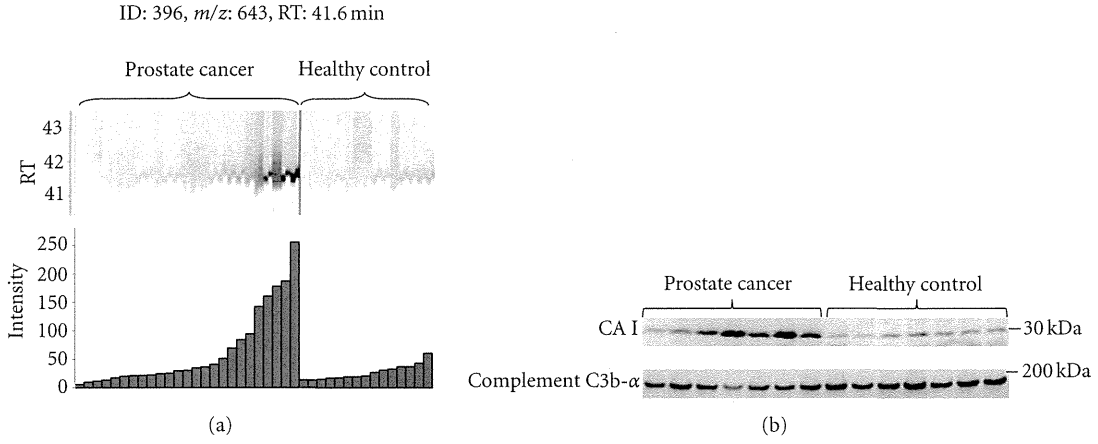

FIGURE 2: (a) MS peaks of peak ID 396 in triplicate LC-MS runs (25 with prostate cancer (left) and 15 with healthy controls (right)) aligned along LC RT. Columns represent the mean intensities of triplicate runs. (b) Immunoblotting of CAI and complement C3b-α (loading control) for 7 prostate cancer patients selected from the most intense MS peaks and 7 healthy controls from the least intense peaks.

Fixed cells were permeabilized with 0.1% Triton in PBS for 5 min, washed 3 times with PBS, blocked for 30 min with 5% normal donkey serum (Chemicon International, Inc., Temecula, CA) in PBS, and stained with the indicated primary antibodies diluted in 1% normal donkey serum overnight at room temperature. After washing 3 times with PBS, fluorophore-conjugated secondary antibodies were applied for 1 h at room temperature and washed with PBS. Alexa Fluor 568-phalloidin was used to visualize the actin cytoskeleton. The slides were mounted with Vectashield antifade reagent (Vector Laboratories), covered with glass coverslips, and observed under Zeiss LSM510 fluorescence microscope equipped with 488/514-nm argon and 543-nm helium-neon lasers.

To stain human tissues, 10 radical prostatectomy specimens were selected with five high and five low plasma CAI concentrations. Sections (4 μm) from 10% formalin-fixed paraffin-embedded material were deparaffinized in xylene and rehydrated in ethanol. Endogenous peroxidase activity was blocked by methanol containing 0.3% hydrogen peroxidase for 30 min. Microwave heating was used for antigen retrieval. The sections were incubated at 4°C overnight with a primary antibody and then incubated with a second antibody for 40 min at room temperature. The reaction products were visualized using diaminobenzidine tetrahydrochloride and counterstained with hematoxylin.

2.9. Statistical Analysis. Statistical significance of intergroup differences was assessed by the Welch's t-test and Mann-Whitney U test. Area under the receiver operating characteristic (ROC) curve (AUC) was calculated for each marker to evaluate its diagnostic significance. A composite index of 2 markers was generated using the results of multivariate logistic regression analysis, which also enabled the calculation of sensitivity, specificity, and ROC curve. Statistical analyses were performed using an open-source statistical language R with the optional module design package.

3. Results

3.1. Plasma Biomarker Discovery by Quantitative MS. To identify a diagnostic biomarker for prostate cancer patients we compared the plasma proteins of 25 prostate cancer patients with those of 15 healthy controls using 2DICAL (Table 1). Among 40,678 independent MS peaks detected within 250–1, 600 m/z, and 30–62.5 min, we found 117 peaks showing significant differences between prostate cancer patients and healthy controls ($P < 0.05$, Welch's t-test). Figure 1(a) shows a representative 2-dimensional view of all MS peaks displayed with m/z along the x-axis and LC retention time along the y-axis. The 117 MS peaks whose expression levels differed significantly between cancer patients and healthy controls are highlighted in red. MS/MS spectra acquired from these 117 MS peaks matched 4 proteins in the database with Mascot scores of >30. Ten peaks matched amino acid sequences of CAI (Table 2). The CAI-derived peak (ID 396) that clearly differed between cancer patients and healthy controls is shown as a representative peak (Figure 1(b)). The intensity distribution of peak ID 396 was different between prostate cancer patients (left) and healthy controls (right) (Figure 2(a)). Immunoblotting with a CAI probe confirmed the results of the 2DICAL findings (Figure 2(b)).

3.2. Large-Scale Validation of Plasma CAI by ELISA. To further validate plasma CAI levels in prostate cancer patients determined using 2DICAL, we performed ELISA to quantify plasma CAI levels in numerous plasma samples. The plasma samples were derived from patients suffering from prostate cancer ($n = 54$), prostatitis ($n = 6$), BPH ($n = 22$), and RCC ($n = 20$) as well as from healthy controls ($n = 60$). Plasma CAI levels were significantly different between prostate cancer patients and healthy controls ($P = 0.022$, Mann-Whitney U test). Plasma CAI levels of patients with BPH or RCC did not show significant differences from healthy

Table 2: Summary of protein identification by tandem mass spectrometry.

ID	*m/z*	RT* (min)	Charge	Normal (mean ± SD)	PCa* (mean ± SD)	*P* Values**	Mascot score	Peptide sequence	Protein description	Uniprot ID
396	643.7	41.6	3	27.35 ± 13.42	63.53 ± 67.39	0.015	95.02	HDTSLKPISVSYNPATAK	Carbonic anhydrase 1	CAH1_HUMAN
310	790.9	47.8	2	24.96 ± 9.73	62.84 ± 72.61	0.016	72.71	ESISVSSEQLAQFR	Carbonic anhydrase 1	CAH1_HUMAN
983	753.0	47.5	3	16.02 ± 4.60	30.40 ± 27.50	0.017	71.62	EIINVGHSFHVNFEDNDNR	Carbonic anhydrase 1	CAH1_HUMAN
1463	719.4	51.8	2	25.20 ± 22.26	11.26 ± 2.32	0.030	66.87	GLEEELQFSLGSK	Complement C4-A	CO4A_HUMAN
1656	872.4	46.0	2	16.99 ± 5.22	25.86 ± 19.63	0.041	66.24	LYPIANGNNQSPVDIK	Carbonic anhydrase 1	CAH1_HUMAN
311	513.8	37.2	2	37.71 ± 16.18	86.33 ± 85.17	0.010	59.09	YSSLAEAASK	Carbonic anhydrase 1	CAH1_HUMAN
798	585.3	45.4	2	21.08 ± 6.68	43.99 ± 34.82	0.004	56.32	SADFTNFDPR	Carbonic anhydrase 2	CAH2_HUMAN
278	485.8	47.6	2	40.70 ± 20.28	91.91 ± 92.78	0.013	50.23	VLDALQAIK	Carbonic anhydrase 1	CAH1_HUMAN
538	493.2	41.0	2	20.19 ± 5.79	47.30 ± 52.18	0.016	44.71	GGPFSDSYR	Carbonic anhydrase 1	CAH1_HUMAN
2429	680.4	47.5	2	27.43 ± 10.68	20.70 ± 3.90	0.032	43.87	LNDLEDALQQAK	Keratin, type II cytoskeletal 1	K2C1_HUMAN
1369	714.4	48.1	3	21.16 ± 5.02	31.96 ± 19.53	0.014	43.57	YDPSLKPLSVSYDQATSLR	Carbonic anhydrase 2	CAH2_HUMAN
916	494.3	48.9	2	12.19 ± 4.64	31.68 ± 32.02	0.006	39.1	VVDVLDSIK	Carbonic anhydrase 2	CAH2_HUMAN
747	920.8	45.3	3	14.00 ± 4.82	35.85 ± 40.39	0.013	36.65	SLLSNVEGDNAVPMQHNNRPTQPLK	Carbonic anhydrase 1	CAH1_HUMAN
1596	656.3	36.6	2	9.93 ± 2.84	19.94 ± 16.30	0.006	33.83	QSPVDIDTHTAK	Carbonic anhydrase 2	CAH2_HUMAN
3781	826.1	40.2	3	4.72 ± 0.94	9.10 ± 9.25	0.027	32.29	TSETKHDTSLKPISVSYNPATAK	Carbonic anhydrase 1	CAH1_HUMAN
1199	538.3	40.7	3	19.72 ± 3.65	30.15 ± 23.16	0.036	30.33	YSAELHVAHWNSAK	Carbonic anhydrase 1	CAH1_HUMAN

*Abbreviation: RT: retention time; PCa: prostate cancer

**Welch's *t*-test.

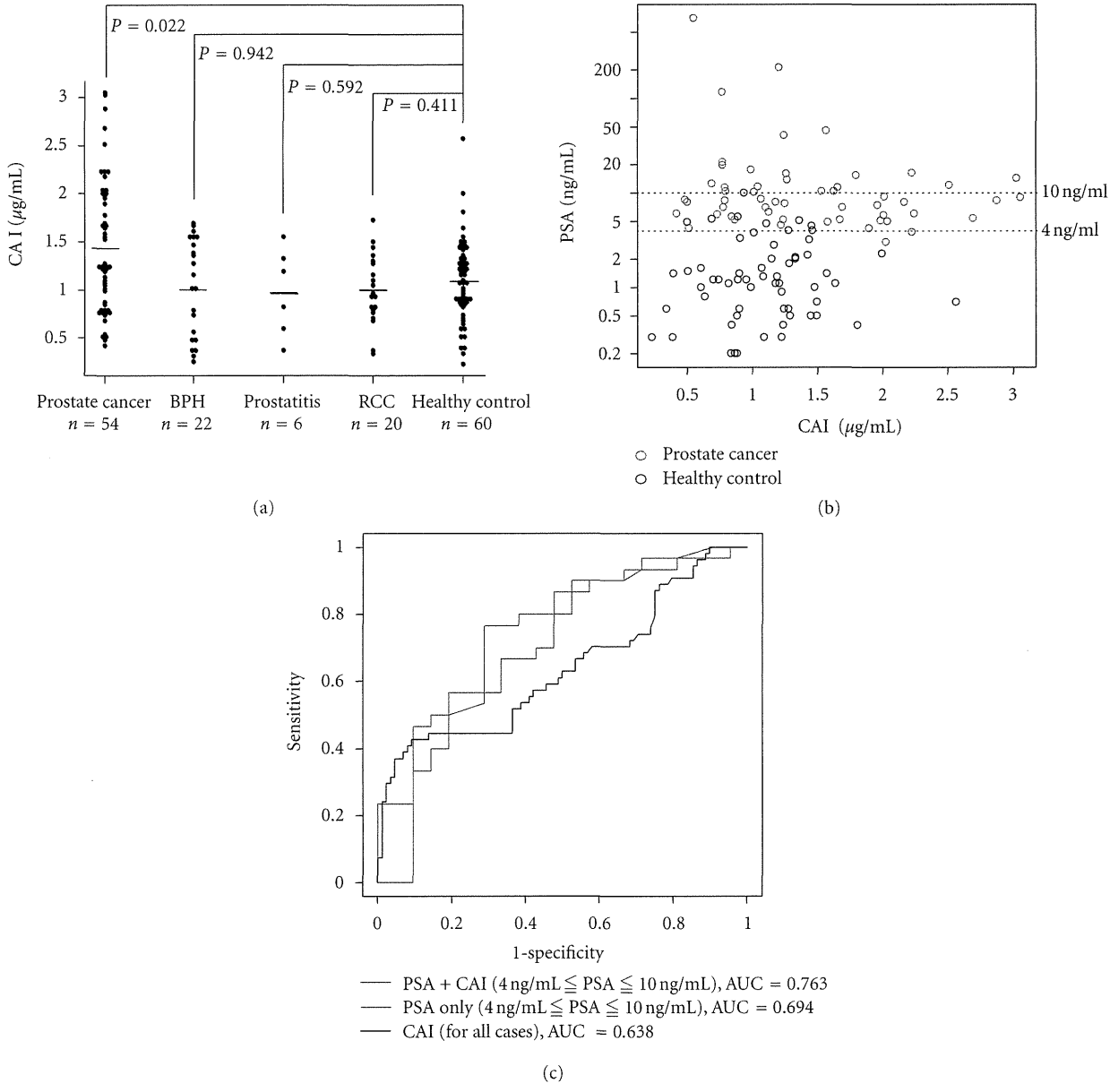

FIGURE 3: (a) Plasma CAI levels in patients with prostate cancer ($n = 54$), BPH ($n = 22$), prostatitis ($n = 6$), and RCC ($n = 20$) and healthy controls ($n = 60$) were 1.43 ± 0.69, 1.03 ± 0.51, 0.98 ± 0.46, 0.99 ± 0.36, and $1.09 \pm 1.82\,\mu$g/mL (mean ± SD), respectively. There was a significant difference between prostate cancer patients and healthy controls ($P = 0.020$, Mann-Whitney U test). Horizontal lines represented the average levels. (b) Scatter plot correlating PSA and CAI levels in prostate cancer patients (red) and healthy controls (black). The dotted lines were added to indicate the PSA gray zone. Serum PSA levels were measured using the Tandem-R kit before the first treatment in each patient. (c) ROC curve of PSA plus CAI (red line) and PSA alone (blue line) confined to the cases with PSA levels in the gray zones. ROC curves were created using a composite index of the 2 markers generated from the results of multivariate logistic regression analysis. As a reference, ROC curves of CAI (black line) for all cases were included.

controls. Plasma CAI levels of prostate cancer patients were clearly higher than those of any other patients (Figure 3(a)).

3.3. Combination of Plasma CAI Levels and PSA Assays. To determine whether plasma CAI levels and PSA assays together would be useful for diagnosing prostate cancer, CAI and PSA levels from the same cases were compared (Figure 3(b)). Pearson's CC between them was −0.106, which meant that these 2 proteins had different vectors in blood concentration. From the distribution data in the plot, CAI levels in prostate cancer patients with PSA levels of >20 ng/mL were low and CAI levels were higher in prostate cancer patients with PSA levels in the gray zone, compared to healthy controls.

To understand the diagnostic significance of CAI levels that coincided with PSA levels in the gray zone, subjects

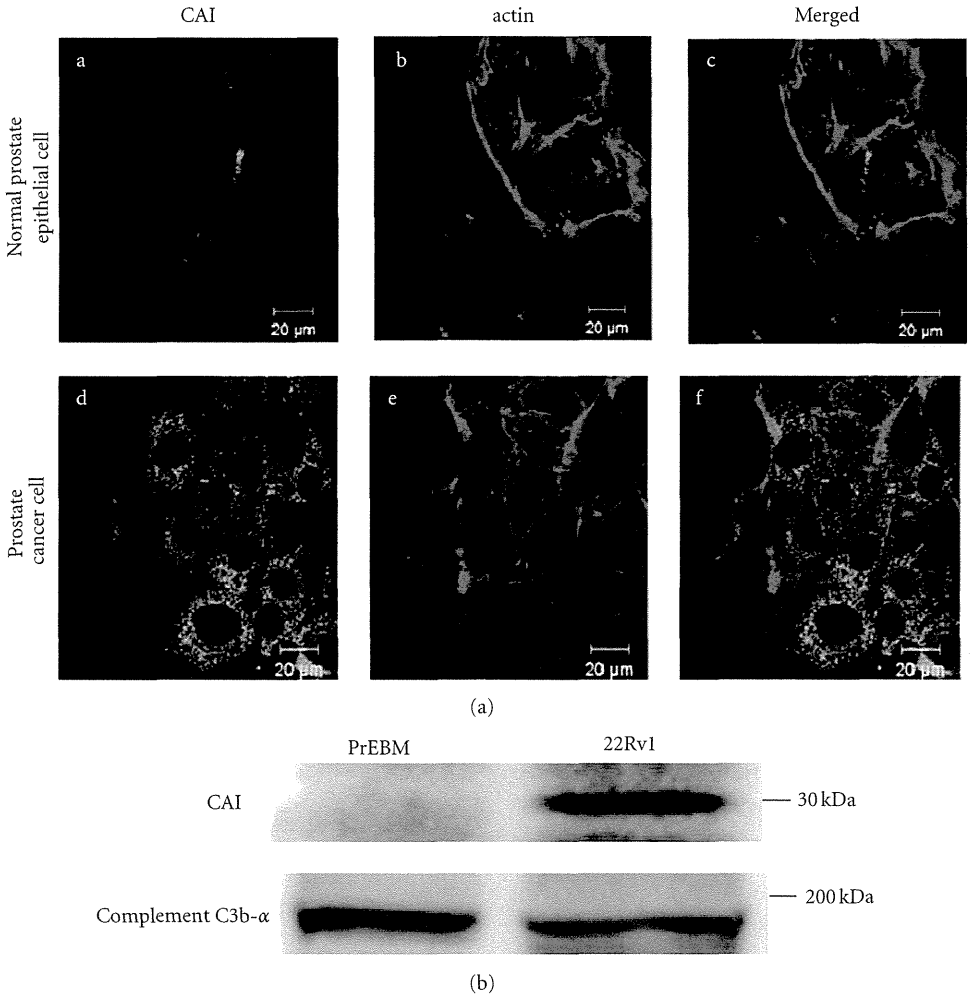

FIGURE 4: CAI expression in cell lines. (a) Immunofluorescence detection of CAI in normal prostate epithelial cells and prostate cancer cells. The cells were fixed for 1 h, permeabilized, and stained with antibodies to detect CAI. Alexa Fluor 488- or Alexa Fluor 568-phalloidin was used to visualize normal prostate epithelial cells (upper panels) and prostate cancer cells of 22Rv1 (lower panels). (b) Western blot analysis of media conditioned by the normal prostate epithelial cells and prostate cancer cells probed using a CAI antibody. C3b-α was used as a loading control.

whose PSA levels were within this gray zone were investigated further. Selected cases were as follows: prostate cancer (n = 30), BPH (n = 10), prostatitis (n = 2), and healthy controls (n = 9). ROC curves for PSA alone and PSA plus CAI were generated for prostate cancer patients relative to the other cases as described in Materials and Methods (Figure 3(c)). AUCs were 0.763 and 0.694, respectively, for PSA plus CAI and PSA alone in the gray zone. PSA levels had a great discriminatory result with an AUC of 0.939 for all cases. However, PSA levels in the gray zone did not provide sufficient discriminatory power when considered alone. This indicates that CAI levels could improve the PSA assay.

3.4. Subcellular Localization and Secretion of CAI. The subcellular localization of CAI was investigated by immunofluorescence microscopy. CAI was hardly detected in the normal prostate epithelial cells. In comparison to actin as a cytoskeletal marker (red, Figure 4(a), b and e), immunofluorescence staining with anti-CAI (green, Figure 4(a), a and d) was observed in the cytoplasm of prostate cancer cells (merged, Figure 4(a), f). CAI was clearly detected in media harvested from 22Rv1 cells (Figure 4(b)).

3.5. The Staining of Human Prostate Cancer Cells. CAI was expressed in every case of prostate cancer. Cancer cells from patients with high plasma CAI concentrations tended to have stronger CAI staining than normal prostatic glands (Figure 5).

4. Discussion

We took advantage of our originally developed label-free proteomic technique (2DICAL) [16] to discover a better biomarkers for prostate cancer diagnosis. We identified that CAI peptide fragments were detected at higher levels in plasma samples from prostate cancer patients than in plasma

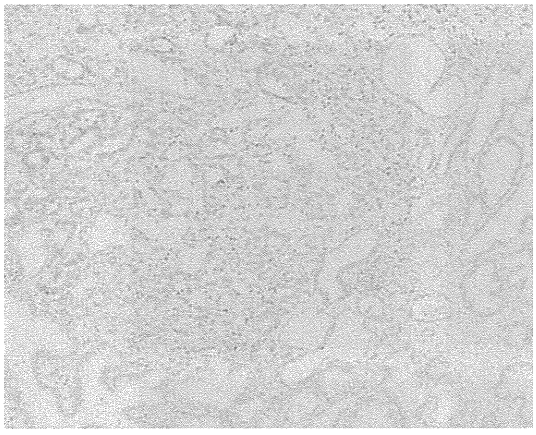

FIGURE 5: Immunohistochemical staining of CAI. CAI was strongly stained at prostate cancer (magnification ×100).

samples from healthy controls. The 2DICAL results were confirmed by Western blot analysis and ELISA using numerous plasma samples including those from patients with urological diseases other than prostate cancer. We found that the combination of CAI and PSA assays has a potential for improving the specificity of PSA assay especially for PSA levels in the gray zone. These initial results suggest that it is reasonable to vigorously pursue CAI as a potentially valuable new biomarker for prostate cancer.

The PSA assay has largely improved the detection of prostate cancer, but only approximately 25% of patients with PSA levels in the gray zone indeed have prostate cancer [19]. Despite the development of several variations of the PSA assay such as free PSA, PSA velocity, or PSA density, these methods do not substantially outperform analysis of total PSA and are not more specific [19, 20]. The reason is because healthy males are known to have PSA levels in the gray zone. Therefore, improved prostate cancer markers such as CAI are eagerly awaited to overcome these problems and enhance diagnostic specificity and sensitivity.

CAI is a zinc metalloenzyme that catalyzes the reversible hydration of carbon dioxide to bicarbonate. Sixteen CA isoforms exist in mammals [21]. Some of them are cytosolic (CAI, CAII, CAIII, CAVII, and CAXIII) catalyzing the hydration of CO_2 to H^+ and HCO_3^-, and subsequently, exporting them from the cell in exchange for Na^+ and Cl^- ions [22]. CAI is a specific marker for the cytoplasm or apical cell membranes of colonic epithelial cells [22, 23] and is related to enterocyte proliferation [24]. CAI is also involved in electroneutral NaCl reabsorption and short chain fatty acid uptake [22]. The biological functions of CAs are of great interest, but CA family members are also being studied as drug targets for treating several diseases such as glaucoma, cancer, obesity, and infections [25].

To the best of our knowledge this is the first report to consider a correlation between CAI blood levels and prostate cancer. To verify the upregulation of plasma CAI in prostate cancer, we investigated the cell biology of prostate cancer using immunofluorescent staining to elucidate its subcellular localization, and Western blotting of culture media to determine if CAI was secreted. We demonstrated increased CAI production and secretion in prostate cancer cell lines. Stronger staining of CAI was observed in prostate cancer cells from patients with high plasma CAI concentrations compared to that in normal prostate glands. However, to clarify the mechanism of the plasma concentration change of CAI, further investigations will be needed considering the secretion mechanism of CAI from prostate cancer cells.

Our present proteome research is new in that we found increased plasma CAI levels in prostate cancer patients. Furthermore, it indicates the possibility that CAI was produced and secreted by prostate cancer cell lines. This study may lead to clinically using CAI as a new prostate cancer marker and the combination of PSA and CAI may have great advantages for diagnosing prostate cancer in patients with gray-zone PSA levels.

Conflict of Interests

The authors have no conflict of interests.

Authors' Contribution

M. Takakura and A. Yokomizo contributed equally to the study.

Acknowledgments

The authors thank Ms. Ayako Ikarashi, Ms. Tomoko Umaki, and Ms. Yuka Nakamura for their technical assistance. Funding was provided by the Program for Promotion of Fundamental Studies in Health Sciences conducted by the National Institute of Biomedical Innovation of Japan, the Third-Term Comprehensive Control Research for Cancer and Research on Biological Markers for New Drug Development conducted by the Ministry of Health and Labor of Japan. These sponsors had no role in the design of the study, collection of the data, analysis and interpretation of the data, decision to submit the paper for publication, or writing of the paper.

References

[1] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, and M. J. Thun, "Cancer statistics, 2009," *CA—A Cancer Journal for Clinicians*, vol. 59, no. 4, pp. 225–249, 2009.

[2] W. J. Catalona, D. S. Smith, T. L. Ratliff, and J. W. Basler, "Detection of organ-confined prostate cancer is increased through prostate- specific antigen-based screening," *JAMA*, vol. 270, no. 8, pp. 948–954, 1993.

[3] J. Pannek and A. W. Partin, "Prostate-specific antigen: what's new in 1997," *Oncology*, vol. 11, no. 9, pp. 1273–1278, 1997.

[4] S. Loeb, S. N. Gashti, and W. J. Catalona, "Exclusion of inflammation in the differential diagnosis of an elevated prostate-specific antigen (PSA)," *Urologic Oncology*, vol. 27, no. 1, pp. 64–66, 2009.

[5] W. J. Catalona, J. P. Richie, F. R. Ahmann et al., "Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: Results of a

multicenter clinical trial of 6,630 men," *Journal of Urology*, vol. 151, no. 5, pp. 1283–1290, 1994.

[6] W. J. Catalona, D. S. Smith, and D. K. Ornstein, "Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/mL and benign prostate examination: Enhancement of specificity with free PSA measurements," *JAMA*, vol. 277, no. 18, pp. 1452–1455, 1997.

[7] A. Magklara, A. Scorilas, W. J. Catalona, and E. P. Diamandis, "The combination of human glandular Kallikrein and free prostrate- specific antigen (PSA) enhances discrimination between prostate cancer and benign prostatic hyperplasia in patients with moderately increased total PSA," *Clinical Chemistry*, vol. 45, no. 11, pp. 1960–1966, 1999.

[8] J. Cervera Deval, F. J. Morales Olaya, J. Jornet Fayos, and M. González Añón, "Diagnostic value of the second prostate biopsies in males of risk. Study stratified by value of PSA," *Actas Urologicas Espanolas*, vol. 28, no. 9, pp. 666–671, 2004.

[9] M. Raber, V. Scattoni, A. Salonia, P. Consonni, and P. Rigatti, "Repeated ultrasound-guided transrectal prostate biopsy in patients with negative histologic test," *Archivio Italiano di Urologia, Andrologia*, vol. 72, no. 4, pp. 197–199, 2000.

[10] M. Ono, M. Shitashige, K. Honda et al., "Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectromery," *Molecular and Cellular Proteomics*, vol. 5, no. 7, pp. 1338–1347, 2006.

[11] A. Negishi, M. Ono, Y. Handa et al., "Large-scale quantitative clinical proteomics by label-free liquid chromatography and mass spectrometry," *Cancer Science*, vol. 100, no. 3, pp. 514–519, 2009.

[12] M. Ono, J. Matsubara, K. Honda et al., "Prolyl 4-hydroxylation of α-fibrinogen. A novel protein modification revealed by plasma proteomics," *The Journal of Biological Chemistry*, vol. 284, no. 42, pp. 29041–29049, 2009.

[13] Y. Murakoshi, K. Honda, S. Sasazuki et al., "Plasma biomarker discovery and validation for colorectal cancer by quantitative shotgun mass spectrometry and protein microarray," *Cancer Science*, vol. 102, no. 3, pp. 630–638, 2011.

[14] J. Matsubara, K. Honda, M. Ono et al., "Reduced plasma level of CXC chemokine ligand 7 in patients with pancreatic cancer," *Cancer Epidemiology Biomarkers and Prevention*, vol. 20, no. 1, pp. 160–171, 2011.

[15] A. Yokomizo, M. Takakura, Y. Kanai et al., "Use of quantitative shotgun proteomics to identify fibronectin 1 as a potential plasma biomarker for clear cell carcinoma of the kidney," *Cancer Biomarkers*, vol. 10, no. 3-4, pp. 175–183, 2011.

[16] M. Ono, M. Kamita, Y. Murakoshi et al., "Biomarker discovery of pancreatic and gastrointestinal cancer by 2DICAL: 2-dimensional image-converted analysis of liquid chromatography and mass spectrometry," *International Journal of Proteomics*, vol. 2012, Article ID 897412, 10 pages, 2012.

[17] Y. Tanaka, H. Akiyama, T. Kuroda et al., "A novel approach and protocol for discovering extremely low-abundance proteins in serum," *Proteomics*, vol. 6, no. 17, pp. 4845–4855, 2006.

[18] M. Idogawa, T. Yamada, K. Honda, S. Sato, K. Imai, and S. Hirohashi, "Poly(ADP-ribose) polymerase-1 is a component of the oncogenic T-cell factor-4/beta;-catenin complex," *Gastroenterology*, vol. 128, no. 7, pp. 1919–1936, 2005.

[19] H. Lilja, D. Ulmert, and A. J. Vickers, "Prostate-specific antigen and prostate cancer: Prediction, detection and monitoring," *Nature Reviews Cancer*, vol. 8, no. 4, pp. 268–278, 2008.

[20] S. Loeb and W. J. Catalona, "What to do with an abnormal PSA test," *Oncologist*, vol. 13, no. 3, pp. 299–305, 2008.

[21] C. T. Supuran and A. Scozzafava, "Carbonic anhydrases as targets for medicinal chemistry," *Bioorganic and Medicinal Chemistry*, vol. 15, no. 13, pp. 4336–4350, 2007.

[22] E. R. Swenson, "Distribution and functions of carbonic anhydrase in the gastrointestinal tract," in *Carbonic Anhydrases: Cellular Physiology and Molecular Genetics*, S. J. Dodgson, R. E. Tashian, G. Gros, and N. D. Carter, Eds., pp. 265–287, Plenum Press, New York, NY, USA, 1991.

[23] G. Lonnerholm and P. Wistrand, "Carbonic anhydrase in the human fetal gastrointestinal tract," *Biology of the Neonate*, vol. 44, no. 3, pp. 166–176, 1983.

[24] I. B. Renes, M. Verburg, D. J. P. M. Van Nispen et al., "Epithelial proliferation, cell death, and gene expression in experimental colitis: Alterations in carbonic anhydrase I, mucin MUC2, and trefoil factor 3 expression," *International Journal of Colorectal Disease*, vol. 17, no. 5, pp. 317–326, 2002.

[25] C. T. Supuran, "Carbonic anhydrase inhibition/activation: Trip of a scientist around the world in the search of novel chemotypes and drug targets," *Current Pharmaceutical Design*, vol. 16, no. 29, pp. 3233–3245, 2010.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2012, Article ID 897412, 10 pages
doi:10.1155/2012/897412

Review Article

Biomarker Discovery of Pancreatic and Gastrointestinal Cancer by 2DICAL: 2-Dimensional Image-Converted Analysis of Liquid Chromatography and Mass Spectrometry

Masaya Ono,[1] Masahiro Kamita,[1] Yusuke Murakoshi,[2] Junichi Matsubara,[3] Kazufumi Honda,[1] Banno Miho,[4] Tomohiro Sakuma,[4] and Tesshi Yamada[1]

[1] *Division of Chemotherapy and Clinical Research, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan*
[2] *Third Department of Surgery, Tokyo Medical University, 6-7-1 Nishi-shinjuku, Shinjuku-ku, Tokyo 160-0023, Japan*
[3] *Institute for Stem Cell Biology and Regenerative Medicine (ISCBRM) and Stanford Cancer Center, Stanford University, 265 Campus Drive, Room G2015, Stanford, CA 94305, USA*
[4] *Bio Science Department, Research and Development Center, Mitsui Knowledge Industry Co., Ltd., 2-7-14 Higashinakano, Nakano-Ku, Tokyo 164-8555, Japan*

Correspondence should be addressed to Masaya Ono, masono@ncc.go.jp

Received 2 March 2012; Accepted 18 April 2012

Academic Editor: Fumio Nomura

Biomarkers tested by blood sample are of great use to clinicians as they provide useful information to aid an early and accurate diagnosis. Comprehensive "omics" studies are expected to facilitate the identification of such new biomarkers, and much research is being performed in this area. Our proteomics analysis system of 2-dimensional image-converted analysis of liquid chromatography and mass spectrometry (2DICAL) has successfully identified several new blood biomarkers from the clinical blood samples of pancreatic and colorectal cancer patients.

1. Introduction

Proteomic studies are powerful tools for identifying useful new biomarkers, and much research is currently being performed in this area. However, the blood proteome is extraordinary difficult to analyze because protein concentrations can vary by 12 orders of magnitude [1]. Thus, biomarker discovery using proteomics requires the development of effective pretreatment protocols to reduce the complexity of blood samples. The identification of biomarkers from clinical samples generally needs large numbers of samples to be compared. The same is true for the identification of biomarkers by mass-spectrometry-coupled proteomics [2, 3]. Our proteomics analysis system of 2-dimensional image-converted analysis of liquid chromatography and mass spectrometry (LC/MS; 2DICAL) and the procedure for reducing blood sample complexity have overcome these problems. We report the successful discovery of several new blood biomarkers for pancreatic and colorectal cancer [4, 5].

2. Biomarker Detection

2.1. Recruitment of Clinical Samples. For biomarker discovery, it is important to collect quality-controlled blood samples. We developed a multi-institutional protocol to preserve blood condition during sampling, storing, freezing, and thawing; all samples were collected and managed at the National Cancer Center Research Institute [6].

2.2. Sample Preparation. As the concentrations of different blood proteins can vary by over 12-orders of magnitude, it is essential to remove abundant proteins or to concentrate specific proteins before proteomics analysis. For this purpose, we used lectin affinity column [5, 7], major protein removal column [8], and hollow fiber membrane (HFM) [9, 10].

2.3. Biomarker Discovery. To identify candidate biomarkers from the proteomics data, we utilized our 2DICAL analysis

system that performs a quantitative comparison of unlabeled shotgun proteomics data generated by LC/MS and enables biomarker discovery from a large number of clinical samples. For selecting blood biomarkers, several decades of samples from cancer patients and healthy controls were analyzed by 2DICAL.

2.4. Biomarker Verification. Biomarkers selected by 2DICAL must be verified. As a rule, we first confirmed 2DICAL results using specific antibodies in small-scale immunoblotting assays. Once marker expression was detected and differences in the expression between patient and control samples were confirmed, large-scale verification was conducted. For this purpose, we used in-house reverse phase protein microarrays (RPPA), which can simultaneously assess hundreds of blood samples by antibody staining [11, 12]. Validation at the hundred-sample scale by multiple reaction monitoring/selective reaction monitoring (MRM/SRM) [13] is also ongoing.

3. Novel Applications of Analysis

We developed our original application for biomarker identification, that is, 2DICAL and RPPA.

3.1. 2DICAL. 2DICAL was developed as a shotgun proteomics analysis system. It analyzes the data of mass to charge ratio (m/z), peak intensity, retention time (RT), and each sample generated by LC/MS as the elemental data; it deploys various 2-dimensional images with different combinations of axes using these four elements. From the m/z-RT image, peaks derived from the same peptide in the direction of acquiring time are integrated. By adding algorisms to ensure reproducibility of m/z and RT, the same peak can be compared precisely across different samples, and a statistical comparison of identical peaks in different samples leads to the discovery of specific differentially expressed peptide peaks. Specific peaks are designated by their m/z and RT coordinates, and further analysis is based on these identifiers. Isotopic labeling is not necessary, and large numbers of samples can be analyzed in this way [4, 8].

3.2. RPPA. RPPA is an emerging high-throughput proteomics technique for validating new biomarkers [14, 15]. Furthermore, RPPA requires significantly lower amounts of clinical samples for quantification than established clinical tests such as enzyme-linked immunosorbent assay (ELISA). We made in-house RPPA using ProteoChip glass slides (Proteogen, Seoul, Republic of Korea) to test hundreds of blood samples simultaneously. For this technique, serially diluted samples are randomly plotted in quadruplicate in a 6,144-spot/slide format using a robot. The spotted slides are incubated with the primary antibody and biotinylated secondary antibody and then processed with a streptavidin-horseradish peroxidase conjugate. The stained slides are scanned on a microarray scanner. Statistical evaluation of the fluorescence intensity of individual samples is performed for large-scale validation of biomarker candidates [11, 12].

4. Biomarkers for Pancreatic Cancer and Gastrointestinal Cancer

Several blood biomarkers have already been discovered. Sample recruitment, sample preparation, biomarker discovery, and validation have been described for each biomarker.

4.1. Biomarkers for Pancreatic Cancer. Prolyl-hydroxylated α-fibrinogen [5] and CXC chemokine ligand 7 (CXCL-7) [10] were identified as pancreatic cancer biomarkers.

4.1.1. Prolyl-Hydroxylated α-Fibrinogen

Objective. Screening for pancreatic cancer.

Samples. In total, 86 plasma samples (collected from 43 patients with pancreatic ductal adenocarcinomas and 43 healthy controls) were used for biomarker identification, and 273 plasma samples (collected from 160 patients with pancreatic ductal adenocarcinomas and 113 healthy controls) were used for validation.

Sample Preparation. Samples were treated with concanavalin A (Con A) to reduce plasma protein complexity.

Biomarker Discovery. Samples were subjected to LC/MS and analyzed by 2DICAL. A total of 115325 peaks were detected, and 6 peaks of 412 m/z (RT 13.7 min), 546 m/z (8.3 min), 552 m/z (8.3 min), 827 m/z (8.3 min), 1141 m/z (29.0 min), and 1185 m/z (9.2 min) were statistically significant with >2 fold difference and $P < 0.0005$ (Mann-Whitney U test) between the pancreatic cancer patient group and healthy control group. Three of the 6 peaks were identified as hydroxyproline-modified α-fibrinogen fragments (Figure 1(a)).

Biomarker Validation. An antibody recognizing α-fibrinogen fragments with an ESSSHH P*GIAEFPSR (P*, 4-hydroxyproline) modification was generated and used for small-scale confirmation of the expression of prolyl-hydroxylated α-fibrinogen and the differences in the expression of modified protein between samples of pancreatic cancer patients and healthy controls (Figure 1(b)). A competitive ELISA was developed using this antibody to quantify plasma levels of prolyl-hydroxylated α-fibrinogen. A significant difference in prolyl-hydroxylated α-fibrinogen expression between plasma samples from pancreatic cancer patients and healthy controls was observed ($P = 3.80 \times 10^{-15}$, Mann-Whitney U test; Figure 1(c)).

4.1.2. CXCL-7

Objective. Screening for pancreatic cancer.

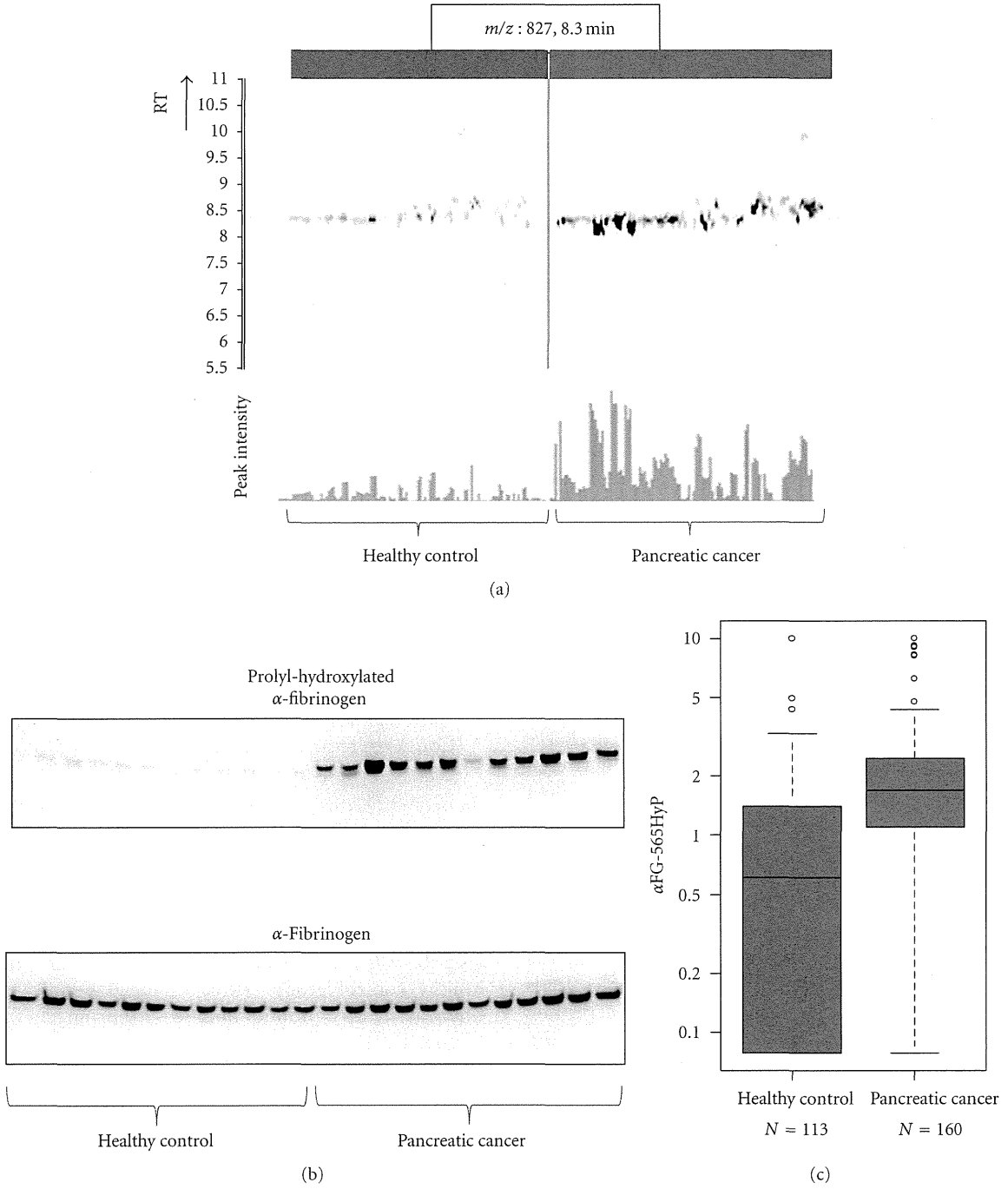

FIGURE 1: Discovery and validation of prolyl-hydroxylated α-fibrinogen as a pancreatic cancer biomarker (partially changed from [5]). (a) 2DICAL images of the peak (*m/z*, 827; RT, 8.3 min) with coordinates RT versus patients (upper) and intensity versus patients (lower). Red indicates samples from pancreatic cancer patients, and blue indicates samples from healthy controls. (b) Western blot of prolyl-hydroxylated α-fibrinogen (upper panel) and total α-fibrinogen (lower panel). (c) Large-scale ELISA validation of the plasma level of prolyl-hydroxylated α-fibrinogen using hundreds of clinical samples.

Samples. A total of 45 plasma samples (collected from 24 patients with pancreatic ductal adenocarcinomas and 21 healthy controls) were used for biomarker discovery and 227 plasma samples (collected from 140 patients with pancreatic ductal adenocarcinomas and 87 healthy controls) were used for biomarker validation.

Sample Preparation. Samples were treated with HFM to reduce plasma protein complexity.

Biomarker Discovery. Samples were subjected to LC/MS and analyzed by 2DICAL. A total of 53009 peaks were detected, and 140 peaks were differentially expressed between

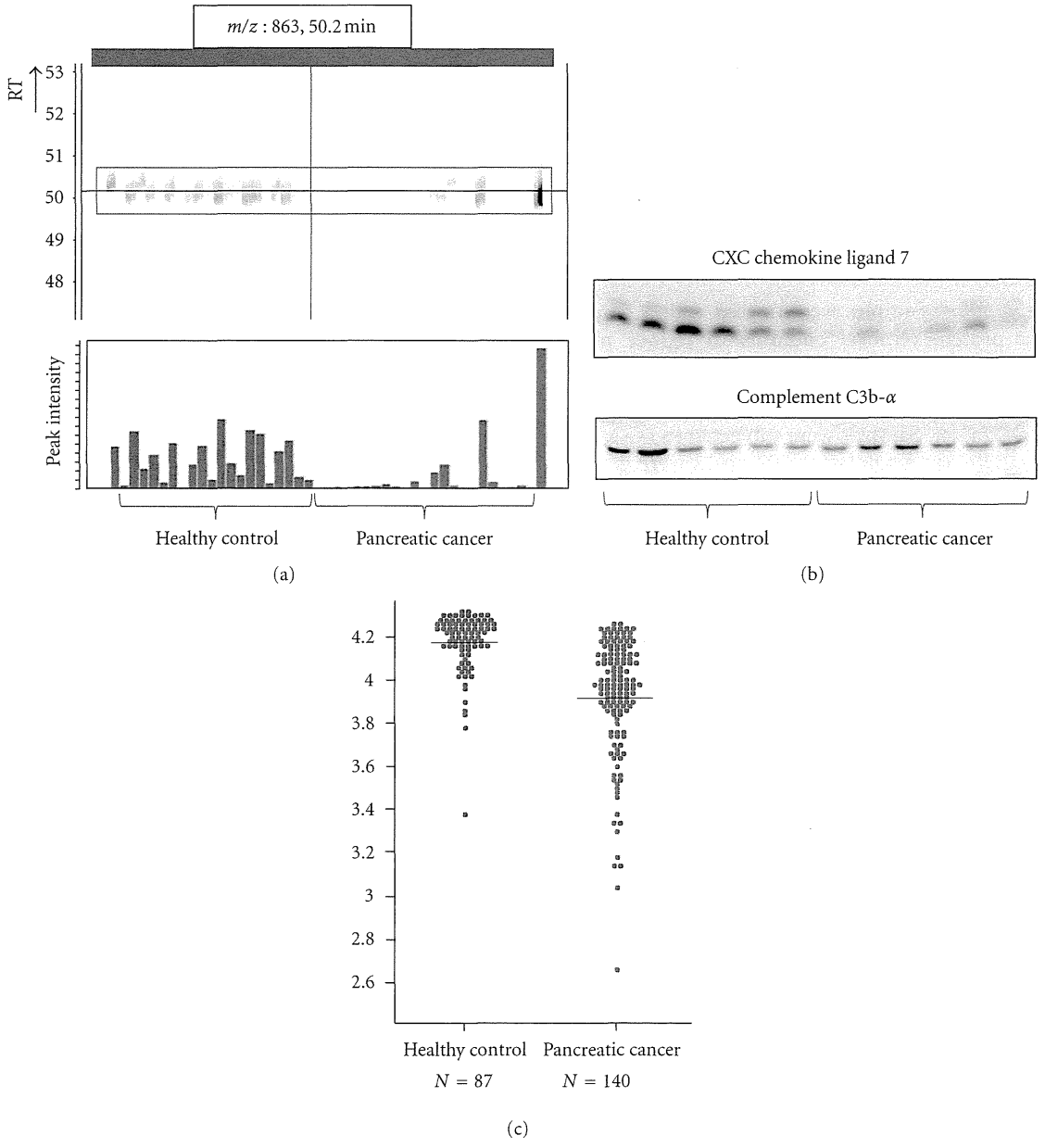

FIGURE 2: Discovery and validation of CXC chemokine ligand 7 as a pancreatic cancer biomarker (partially changed from [10]). (a) 2DICAL images of peptide peak (*m/z*, 863; RT, 50.2 min) with coordinates RT versus patients (upper) and intensity versus patients (bottom). Red indicates samples from pancreatic cancer patients, and blue indicates samples from healthy controls. (b) Western blot of CXC chemokine ligand 7 (upper panel) and the loading control Complement C3b-α. (c) Large-scale RPPA validation of the plasma level of CXC chemokine ligand 7 using hundreds of clinical samples.

pancreatic cancer patients and healthy controls, with an area under curve (AUC) of >0.800. Of these, 10 proteins were annotated by database search of tandem mass spectra. The 862 *m/z* (RT 50.2 min) peak annotated as a fragment of CXCL-7 was specifically expressed in pancreatic cancer patients, with an AUC of 0.839 ($P = 4.54 \times 10^{-5}$ by Mann-Whitney U test) (Figure 2(a)).

Biomarker Validation. Small-scale confirmation of CXCL7 identification and differential expression was done by immunoblotting using an anti-CXCL-7 antibody (Figure 2(b)). For large-scale validation, 227 plasma samples were randomly plotted onto ProteoChip glass slides for RPPA and blotted with an anti-CXCL-7 antibody. CXCL7 expression in pancreatic cancer patients and healthy controls was confirmed to be significantly different ($P = 1.40 \times 10^{-16}$, Welch t-test; Figure 2(c)).

4.2. Biomarkers for Colorectal Cancer. Complement Component 9 (C9) [12] and adipophilin [16] were identified as colorectal cancer biomarkers.

4.2.1. C9

Objective. Screening for colorectal cancer.

Samples. In total, 90 plasma samples (collected from 31 colorectal cancer patients and 59 healthy controls) were used for biomarker discovery, and 345 plasma samples (collected from 115 colorectal cancer patients and 230 healthy controls) were used for validation.

Sample Preparation. Samples were treated with a 12-abundant-plasma-protein removal columns to reduce plasma protein complexity.

Biomarker Discovery. Samples were subjected to LC/MS and analyzed by 2DICAL. A total of 94803 peaks were detected, and 90 peaks showed statistically significant differences in expression between plasma from colorectal cancer patients and healthy controls. Of these, 10 proteins were annotated by database search of tandem mass spectra. A peptide peak with 622 m/z (RT 56.8 min) was annotated as a fragment of C9 specific to colorectal cancer patients ($P = 3.0 \times 10^{-5}$, paired t-test; Figure 3(a)).

Biomarker Validation. Small-scale confirmation of C9 identification and differential expression was done by immunoblotting using an anti-C9 antibody (Figure 3(b)). For large-scale validation, 345 plasma samples were randomly plotted into ProteoChip glass slides for RPPA and blotted with an anti-C9 antibody. There was a significant difference in C9 expression in plasma from colorectal cancer patients and from healthy controls ($P = 1.43 \times 10^{-12}$, Student's t-test; Figure 3(c)).

4.2.2. Adipophilin

Objective. Screening for colorectal cancer.

Samples. A total of 43 plasma samples (collected from 22 colorectal cancer patients and 21 healthy controls) were used for biomarker discovery, and 323 plasma samples (collected from 127 colorectal cancer patients and 196 healthy controls) were used for validation.

Sample Preparation. Samples were treated with HFM to reduce plasma protein complexity.

Biomarker Discovery. Pretreated samples were subjected to LC/MS and analyzed by 2DICAL. A total of 53009 peptide peaks were detected, and 103 peaks with an AUC of >0.800 were differentially expressed in healthy controls and colorectal cancer patients. Of these, 6 proteins were annotated by database search of tandem mass spectra. The 749 m/z (RT 47.4 min) peak represents a fragment of adipophilin specifically present in colorectal cancer patients (0.814 in AUC; Figure 4(a)).

Biomarker Validation. Small-scale confirmation of adipophilin identification and differential expression was done by immunoblotting using an anti-adipophilin antibody (Figure 4(b)). For large-scale validation, 323 plasma samples were randomly plotted into ProteoChip glass slides for RPPA and blotted with an anti-adipophilin antibody. Differential expression of adipophilin between plasma samples from colorectal cancer patients and from healthy controls was significant ($P = 5.49 \times 10^{-10}$, Welch t-test; Figure 4(c)).

4.3. Biomarker for Adverse Effects in Pancreatic Cancer following Chemotherapy

4.3.1. Haptoglobin [17]

Objective. Prediction for the adverse effect of pancreatic cancer chemotherapy.

Samples. A total of 47 plasma samples collected from patients with pancreatic ductal adenocarcinomas and treated with gemcitabine (2′,2′-difluorodeoxycytidine) monotherapy (25 with severe adverse effects (AEs) and 22 without) were used for biomarker discovery, and 253 plasma samples and 52 serum samples were collected from patients with pancreatic ductal adenocarcinomas treated by gemcitabine monotherapy for validation.

Sample Preparation. Samples were treated with a 12 abundant plasma protein removal column to reduce plasma protein complexity.

Biomarker Discovery. Samples were subjected to LC/MS and analyzed by 2DICAL. A total of 60,888 peaks were detected and 757 peaks differed significantly between patients with severe AEs and patients without AEs ($P < 0.001$, Welch t-test). Among these, the peak with highest value to discriminate patients with severe AEs from those without AEs was annotated as haptoglobin. The haptoglobin fragment peak of 491 m/z (RT 44.5 min) is shown in Figure 5(a).

Biomarker Validation. Small-scale confirmation of haptoglobin identification and differential expression was confirmed by immunoblotting using an anti-haptoglobin antibody (Figure 5(b)). Haptoglobin concentration in 305 plasma and serum samples was measured by immunonephelometry. The severity of AE severity inversely correlated with the concentration of haptoglobin (Figure 5(c)).

4.4. Biomarker for Predicting Survival of Pancreatic Cancer Patients following Chemotherapy

4.4.1. α1-Antitrypsin [11]

Objective. Prediction of the survival for pancreatic cancer chemotherapy.

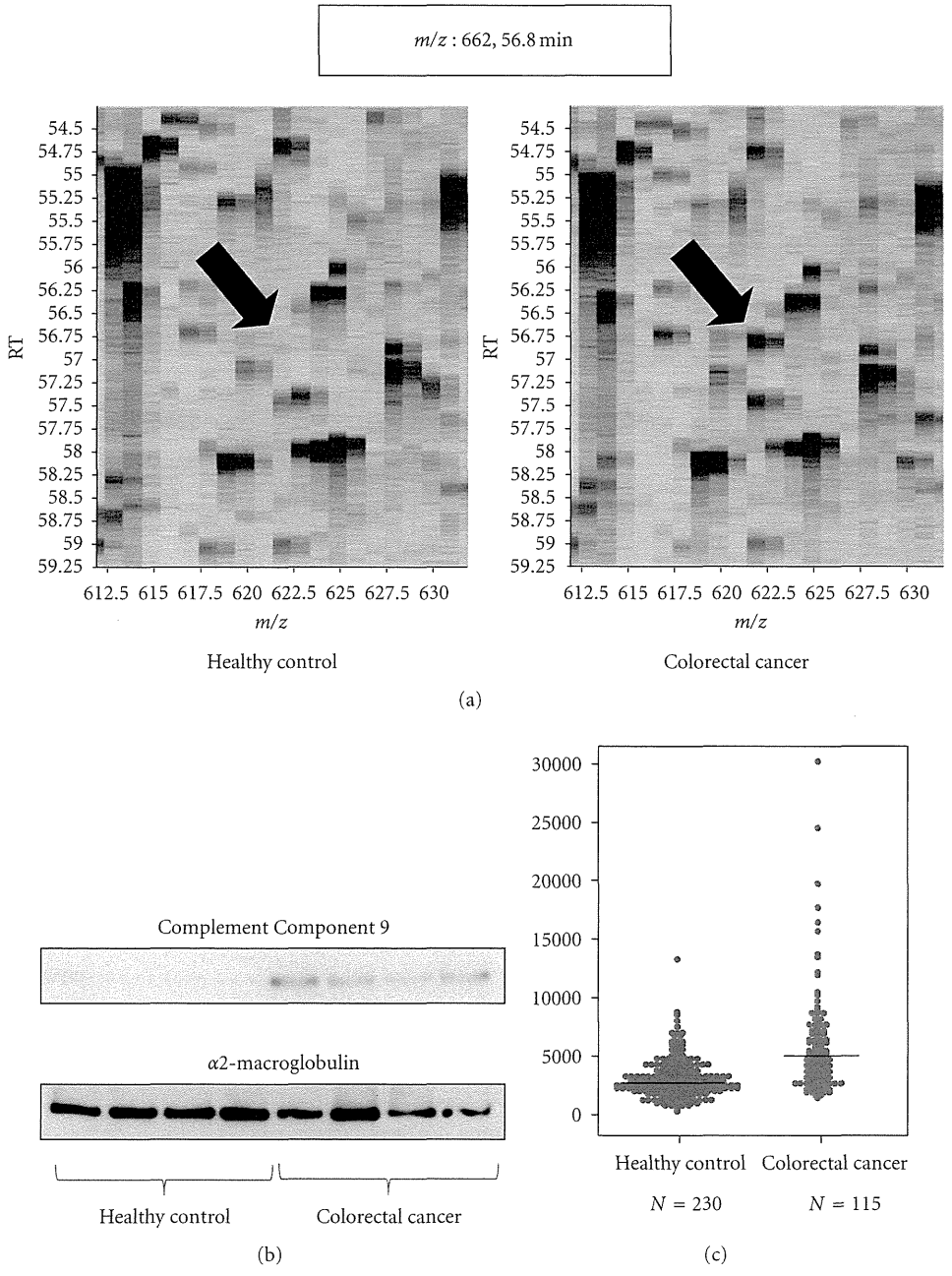

FIGURE 3: Discovery and validation of Complement Component 9 as a colorectal cancer biomarker (partially changed from [12]). (a) 2DICAL images with coordinates *m/z* versus RT. The intensity of the peak of 622 *m/z* and RT of 56.8 min (indicated by arrows) are clearly different in the plasma samples from healthy controls (left) and colorectal cancer patients (right). (b) Western blot of Complement Component 9 and the loading control α2-macroglobulin. (c) Large-scale RPPA validation of the plasma level of Complement Component 9 using hundreds of clinical samples.

Samples. A total of 60 plasma samples collected from patients with pancreatic ductal adenocarcinomas and treated by gemcitabine monotherapy (29 with short-term survival and 31 with long-term survival) were used for biomarker discovery, and 304 samples collected from patients with pancreatic ductal adenocarcinomas and treated by gemcitabine monotherapy were used for validation.

Sample Preparation. Samples were treated with 12-abundant-plasma-protein removal column to reduce plasma protein complexity.

Biomarker Discovery. Samples were subjected to LC/MS and analyzed by 2DICAL. A total of 45227 peaks were detected, and 637 peaks differed significantly between patients with

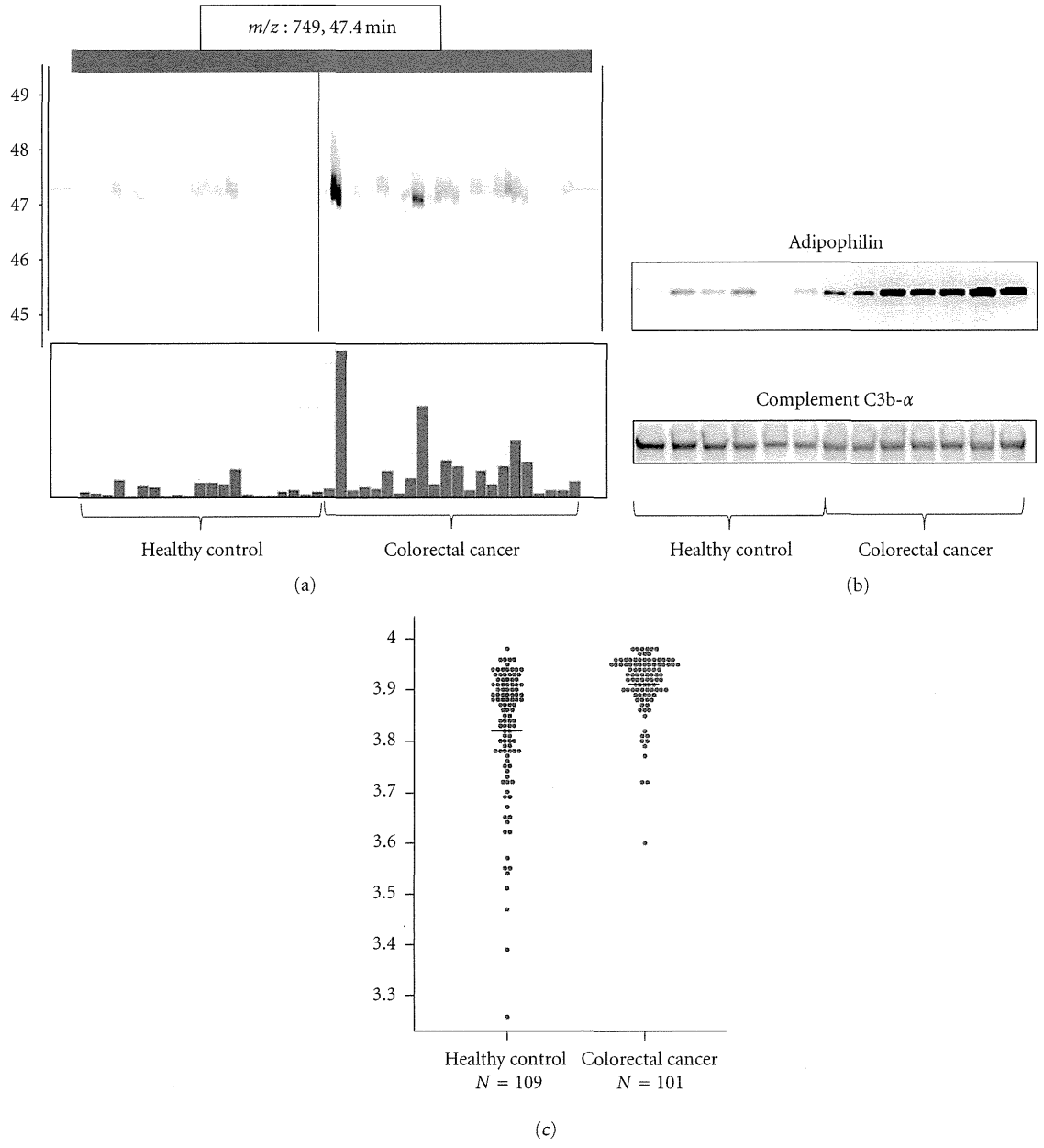

FIGURE 4: Discovery and validation of adipophilin as a colorectal cancer biomarker (partially changed from [16]). (a) 2DICAL images of the peak (*m/z*, 749; RT, 47.4 min) with coordinates RT versus patients (upper) and intensity versus patients (lower). Red indicates samples from colorectal cancer patients, and blue indicates samples from healthy controls. (b) Western blot of adipophilin and the loading control Complement C3b-α. (c) Large-scale RPPA validation of the plasma level of adipophilin using hundreds of clinical samples.

long-term survival and those with short-term survival ($P < 0.001$, Welch t-test). The peptide peak that best discriminated patients with short-term survival from those with long-term survival ($P = 2.57 \times 10^{-4}$) at 491 *m/z* (RT 44.5 min) was annotated as a fragment of a1-antitrypsin (Figure 6(a)).

Biomarker Validation. Small-scale confirmation of α1-antitrypsin identification and differential expression was done by immunoblotting using an anti-α1-antitrypsinantibody (Figure 6(b)). For large-scale validation, 304 samples were randomly plotted into ProteoChip glass slides for RPPA and blotted with antibody to α1-antitrypsin. Improved survival of patients with pancreatic ductal adenocarcinoma treated by gemcitabine monotherapy correlated with low blood concentrations of α1-antitrypsin (Figure 6(c)).

5. Conclusions

We have established a comprehensive method for identifying blood biomarkers, which covers all aspects of analysis from

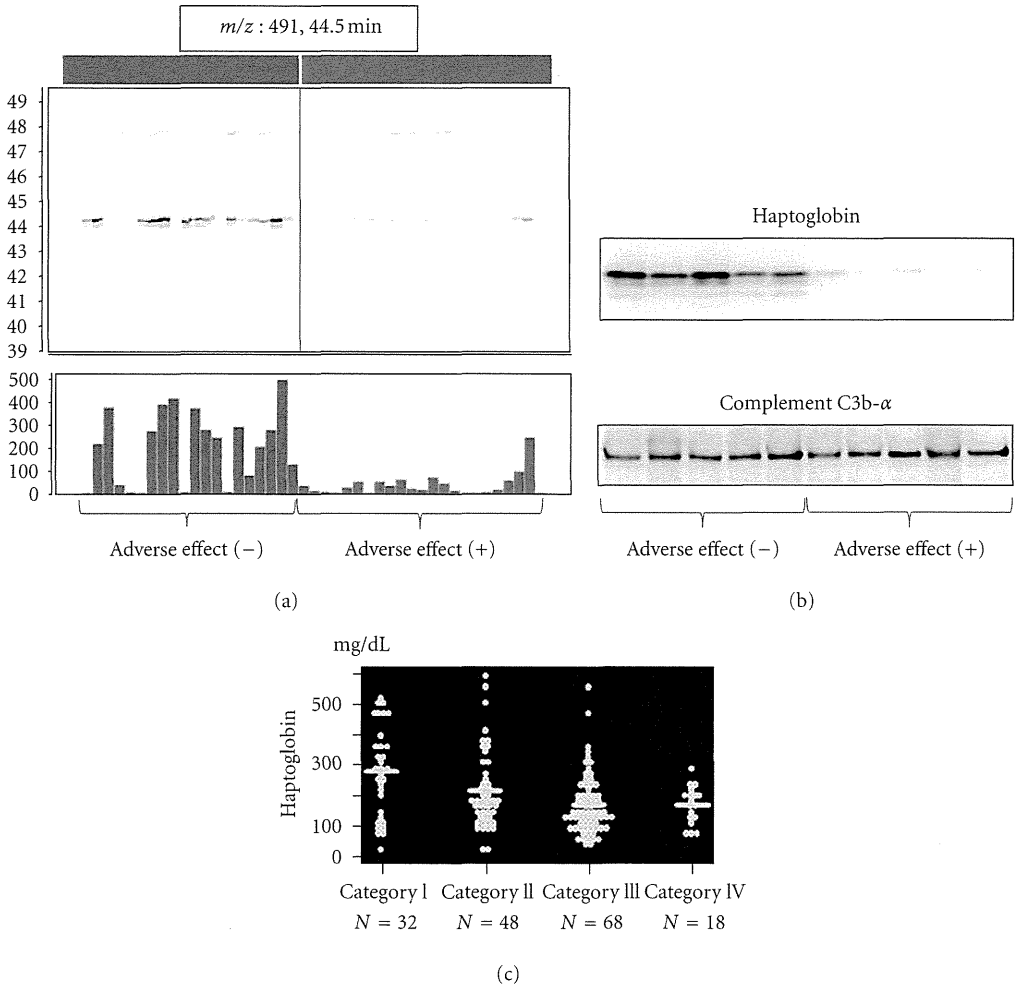

FIGURE 5: Discovery and validation of haptoglobin as a biomarker for adverse effects in pancreatic cancer following chemotherapy (partially changed from [17]). (a) 2DICAL images of the peak (m/z, 409; RT, 44.5 min) with coordinates RT versus patients (upper) and intensity versus patients (lower). Red indicates samples from pancreatic cancer patients with severe AEs following chemotherapy, and blue indicates samples from pancreatic cancer patients without AEs following chemotherapy. (b) Western blot of haptoglobin and the loading control Complement C3b-α. (c) Large-scale immunonephelometric validation of the plasma level of haptoglobin using hundreds of clinical samples. The adverse effects were categorized in four grades according to the degree of neutropenia and thrombocytopenia. Haptoglobin concentration decreased according to the increase of the adverse effect severity.

sample recruitment to biomarker discovery and validation. The next stage in the development of these novel biomarkers is to test them in a clinical context. The proteomics approach for blood biomarker discovery identifies a new function for common proteins such as these biomarkers. With technological advances in sample preparations, resolution and sensitivity of mass spectrometer, and methods for the identification of proteins from mass spectra, we can expect to discover biomarkers existing in much smaller amount or those with new structures in the future. We also expect that large-scale validation of biomarkers discovered using mass spectrometer will be conducted by MRM/SRM. 2DICAL is applicable not only for proteomics but also for metabolomics or glycomics and has a great potential for identifying disease-associated post-translational protein modifications. 2DICAL will evolve along with technological advances and contribute the discovery of new biomarkers in future.

Acknowledgments

The authors thank Ms. Ayako Ikarashi, Ms. Tomoko Umaki, and Ms. Yuka Nakamura for their technical assistance. Funding was provided by the Program for Promotion of Fundamental Studies in Health Sciences conducted by the National Institute of Biomedical Innovation of Japan, the Third-Term Comprehensive Control Research for Cancer and Research on Biological Markers for New Drug Development conducted by the Ministry of Health and Labor of Japan. These sponsors had no role in the design of the study, collection of the data, analysis and interpretation of the data, decision to submit the paper for publication, or writing of the paper.

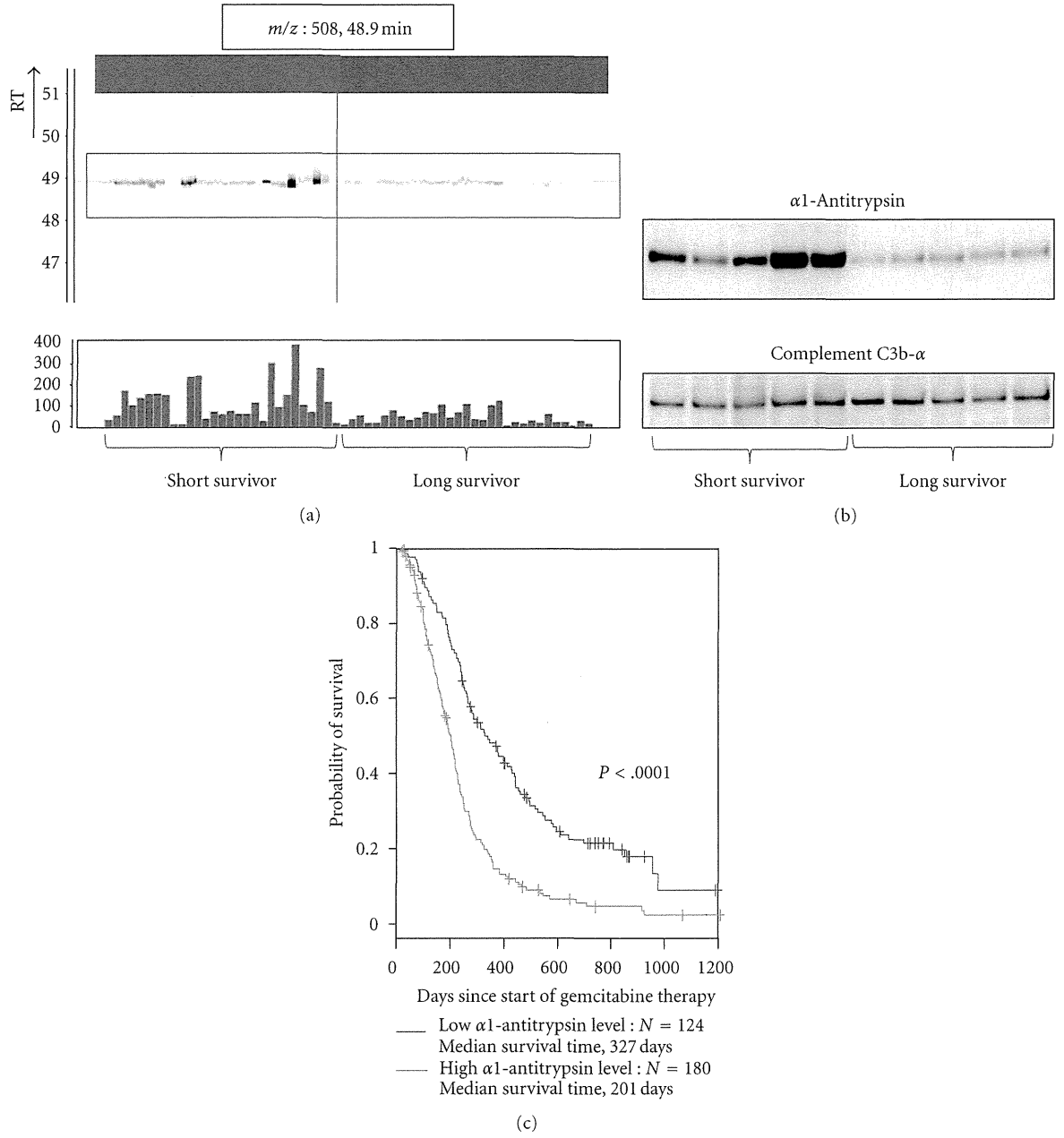

FIGURE 6: Discovery and validation of α1-antitrypsin as a biomarker for predicting survival of pancreatic cancer patients following chemotherapy (partially changed from [11]). (a) 2DICAL images of the peak (m/z, 508; RT, 48.9 min) with coordinates RT versus patients (upper) and intensity versus patients (lower). Red indicates samples from pancreatic cancer patients with short-term survival, and blue indicates samples from pancreatic cancer patients with long-term survival. (b) Western blot of α1-antitrypsin and the loading control Complement C3b-α. (c) Large-scale RPPA validation of the plasma level of α1-antitrypsin using hundreds of clinical samples. Survival curve was significantly better in the group of low α1-antitrypsin level than that of high α1-antitrypsin level.

References

[1] N. L. Anderson and N. G. Anderson, "The human plasma proteome: history, character, and diagnostic prospects," *Molecular & Cellular Proteomics*, vol. 1, no. 11, pp. 845–867, 2002.

[2] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold, "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags," *Nature Biotechnology*, vol. 17, no. 10, pp. 994–999, 1999.

[3] L. DeSouza, G. Diehl, M. J. Rodrigues et al., "Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry," *Journal of Proteome Research*, vol. 4, no. 2, pp. 377–386, 2005.

[4] M. Ono, M. Shitashige, K. Honda et al., "Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectromery," *Molecular and Cellular Proteomics*, vol. 5, no. 7, pp. 1338–1347, 2006.

[5] M. Ono, J. Matsubara, K. Honda et al., "Prolyl 4-hydroxylation of α-fibrinogen. A novel protein modification revealed by

plasma proteomics," *Journal of Biological Chemistry*, vol. 284, no. 42, pp. 29041–29049, 2009.

[6] K. Honda, Y. Hayashida, T. Umaki et al., "Possible detection of pancreatic cancer by plasma protein profiling," *Cancer Research*, vol. 65, no. 22, pp. 10613–10622, 2005.

[7] S. Ogata, T. Muramatsu, and A. Kobata, "Fractionation of glycopeptides by affinity column chromatography on concanavalin A Sepharose," *Journal of Biochemistry*, vol. 78, no. 4, pp. 687–696, 1975.

[8] A. Negishi, M. Ono, Y. Handa et al., "Large-scale quantitative clinical proteomics by label-free liquid chromatography and mass spectrometry," *Cancer Science*, vol. 100, no. 3, pp. 514–519, 2009.

[9] Y. Tanaka, H. Akiyama, T. Kuroda et al., "A novel approach and protocol for discovering extremely low-abundance proteins in serum," *Proteomics*, vol. 6, no. 17, pp. 4845–4855, 2006.

[10] J. Matsubara, K. Honda, M. Ono et al., "Reduced plasma level of CXC chemokine ligand 7 in patients with pancreatic cancer," *Cancer Epidemiology Biomarkers and Prevention*, vol. 20, no. 1, pp. 160–171, 2011.

[11] J. Matsubara, M. Ono, K. Honda et al., "Survival prediction for pancreatic cancer patients receiving gemcitabine treatment," *Molecular and Cellular Proteomics*, vol. 9, no. 4, pp. 695–704, 2010.

[12] Y. Murakoshi, K. Honda, S. Sasazuki et al., "Plasma biomarker discovery and validation for colorectal cancer by quantitative shotgun mass spectrometry and protein microarray," *Cancer Science*, vol. 102, no. 3, pp. 630–638, 2011.

[13] E. M. H. Finlay, D. E. Games, J. R. Startin, and J. Gilbert, "Screening, confirmation, and quantification of sulphonamide residues in pig kidney by tandem mass spectrometry of crude extracts," *Biomedical and Environmental Mass Spectrometry*, vol. 13, no. 11, pp. 633–639, 1986.

[14] T. Grote, D. R. Siwak, H. A. Fritsche et al., "Validation of reverse phase protein array for practical screening of potential biomarkers in serum and plasma: accurate detection of CA19-9 levels in pancreatic cancer," *Proteomics*, vol. 8, no. 15, pp. 3051–3060, 2008.

[15] J. D. Wulfkuhle, L. A. Liotta, and E. F. Petricoin, "Proteomic applications for the early detection of cancer," *Nature Reviews Cancer*, vol. 3, no. 4, pp. 267–275, 2003.

[16] J. Matsubara, K. Honda, M. Ono et al., "Identification of adipophilin as a potential plasma biomarker for colorectal cancer using label-free quantitative mass spectrometry and protein microarray," *Cancer Epidemiology, Biomarkers and Prevention*, vol. 20, no. 10, pp. 2195–2203, 2011.

[17] J. Matsubara, M. Ono, A. Negishi et al., "Identification of a predictive biomarker for hematologic toxicities of gemcitabine," *Journal of Clinical Oncology*, vol. 27, no. 13, pp. 2261–2268, 2009.

Supplemental Material can be found at:
http://www.mcponline.org/cgi/content/full/M900044-MCP200
/DC1

Snapshot Peptidomics of the Regulated Secretory Pathway*[S]

Kazuki Sasaki‡§, Yoshinori Satomi¶‖, Toshifumi Takao¶, and Naoto Minamino‡**

Neurons and endocrine cells have the regulated secretory pathway (RSP) in which precursor proteins undergo proteolytic processing by prohormone convertase (PC) 1/3 or 2 to generate bioactive peptides. Although motifs for PC-mediated processing have been described ((R/K)X_n(R/K) where n = 0, 2, 4, or 6), actual processing sites cannot be predicted from amino acid sequences alone. We hypothesized that discovery of bioactive peptides would be facilitated by experimentally identifying signal peptide cleavage sites and processing sites. However, *in vivo* and *in vitro* peptide degradation, which is widely recognized in peptidomics, often hampers processing site determination. To obtain sequence information about peptides generated in the RSP on a large scale, we applied a brief exocytotic stimulus (2 min) to cultured endocrine cells and analyzed peptides released into supernatant using LC-MSMS. Of note, 387 of the 400 identified peptides arose from 19 precursor proteins known to be processed in the RSP, including nine peptide hormone and neuropeptide precursors, seven granin-like proteins, and three processing enzymes (PC1/3, PC2, and peptidyl-glycine α-amidating monooxygenase). In total, 373 peptides were informative enough to predict processing sites in that they have signal sequence cleavage sites, PC consensus sites, or monobasic cleavage sites. Several monobasic cleavage sites identified here were previously proved to be generated by PCs. Thus, our approach helps to predict processing sites of RSP precursor proteins and will expedite the identification of unknown bioactive peptides hidden in precursor sequences. ***Molecular & Cellular Proteomics 8:1638–1647, 2009.***

The generation of peptide hormones or neuropeptides involves the proteolytic processing of precursor proteins by specific proteases. In neurons and endocrine cells, most, if not all, of these bioactive peptides are generated within the RSP[1] in which the processing enzymes PC1/3 or PC2 cleave precursors at basic residues (1, 2). The PC-mediated cleavage most often occurs at consecutive basic residues, but not all basic residues serve as PC recognition sites (2). This is partly because the secondary structure of a precursor also affects the substrate recognition (3). Identification of processing sites is hence a prerequisite for locating unknown peptides hidden in a precursor sequence.

Peptidomics has been advocated to comprehensively study peptides cleaved off from precursor proteins by endogenous proteases (4–6). These naturally occurring peptides are beyond the reach of current proteomics and should be analyzed in their native forms. Unlike proteomics, peptidomics has the potential to uncover processing sites of precursor proteins. Most peptidomics studies, which target tissue peptidomes from brain or endocrine organs (7–11), have provided limited information about secretory peptides that could help to identify processing sites; they are too often blurred by subsequent actions of exopeptidases (cutting off a single amino acid or dipeptide from either end of a peptide).

In MS-based identification of bioactive peptides present in biological samples, their relative low abundance in a total pool of naturally occurring peptides should be considered. Once extracted from cultured cells or tissues, *bona fide* secretory peptides and nonsecretory peptides or peptide fragments caused by degradation of abundant cytosolic proteins cannot be discriminated, and therefore we need to analyze samples rich in secretory peptides to facilitate the identification of bioactive peptides. Several attempts have been made to isolate secretory proteins or peptides, such as subcellular fractionation for harvesting secretory granules (12, 13). With all these efforts, a limited number of secretory peptides have been identified, and many known bioactive peptides still escape analysis.

We took advantage of the fact that peptides processed in the RSP are enriched in secretory granules of neurons and endocrine cells and released on exocytosis. Here we applied a brief exocytotic stimulus (2 min) to cultured human endocrine cells and identified peptides released into supernatant using LC-MSMS on an LTQ-Orbitrap mass spectrometer. Nearly 97% of the identified peptides arose from precursor proteins known to be recruited to the RSP, such as peptide hormone precursors and granin-like secretory proteins. Our

From the ‡Department of Pharmacology, National Cardiovascular Center Research Institute, Fujishirodai, Suita, Osaka 565-8565, Japan and ¶Laboratory of Protein Profiling and Functional Proteomics, Institute for Protein Research, Osaka University, Yamadaoka, Suita, Osaka 565-0871, Japan

Received, January 27, 2009, and in revised form, March 25, 2009
Published, MCP Papers in Press, March 31, 2009, DOI 10.1074/mcp.M900044-MCP200

[1] The abbreviations used are: RSP, regulated secretory pathway; CgA, chromogranin A; CgB, chromogranin B; CT, calcitonin; CGRP, calcitonin gene-related peptide; GRP, gastrin-releasing peptide; PC, prohormone convertase; SgII, secretogranin II; SgIII, secretogranin III; SST, somatostatin; LTQ, linear trap quadrupole; IPI, International Protein Index.

approach was validated by the identification of previously known processing sites of peptide hormone precursors. In addition, a majority of the identified peptides retained cleavage sites that agree with consensus cleavage sites for PCs, which are informative enough to deduce the processing sites of RSP proteins. This peptidomics approach will expedite the identification of unknown bioactive peptides.

EXPERIMENTAL PROCEDURES

Peptide Preparation—Monolayer cultures of TT cells (14, 15) were rinsed three times with Hanks' medium (Invitrogen). Culture supernatants of the cells incubated for 2 min before and after stimulation with 10 μM forskolin plus 10 μM carbachol were harvested and rapidly extracted at 4 °C using an RP-1 solid phase extraction cartridge (GL Sciences) without centrifuging the supernatants. Bound substances were eluted in 60% ACN, 0.1% formic acid. After lyophilization of a small aliquot, samples were reconstituted in 10 μl of 2% ACN, 0.1% formic acid. In the one-dimensional analysis, solid phase-extracted analytes were subjected to LC-MSMS without gel filtration using an aliquot equivalent to 5×10^5 cells. Peptide fractions were obtained by HPLC on a gel filtration column equilibrated with 60% ACN, 0.1% TFA at a flow rate of 1.5 ml/min (G2000SWXL, 21.5 × 300 mm, Tosoh Corp.). For gel-filtrated fractions, an aliquot corresponding to 1.5×10^6 cells was used for LC-MSMS.

LC-MSMS—Nano-LC-MSMS experiments were performed with a Chorus nanoflow system (CS Analytics) connected to an LTQ-Orbitrap mass spectrometer (ThermoFisher Scientific) equipped with a nanoelectrospray emitter (MonoSpray C_{18} Nano, 100 μm × 50 mm, GL Sciences). Samples were dissolved in solvent A (2% ACN, 0.1% formic acid). The nanoflow system was run at a flow rate of 500 nl/min with a gradient from 5 to 45% solvent B (89% ACN, 0.1% formic acid) in 40 min and then to 95% B in 1 min. A protonated ion of polycyclodimethylsiloxane with *m/z* 445.120025 was used for internal calibration throughout. The mass spectrometer was operated in a data-dependent mode to automatically switch between MS and MSMS acquisitions. Survey full-scan spectra were acquired in the *m/z* range 400–1500 with five most intense ions (intensity threshold, 2e+05) sequentially isolated for MSMS in the linear ion trap using collision-induced dissociation with dynamic exclusion onward throughout the following scans. The resultant product ions were recorded in the Orbitrap.

Data Analysis and Peptide Identification—Peak picking, deisotoping, and deconvolution of MSMS spectra were preformed using Mascot Distiller (version 2.1.1.0) with the default parameters for Orbitrap. Peak lists were searched against IPI human (72,079 entries on July 2, 2008) using Mascot (version 2.2) with no enzyme specification. Pyroglutamination, C-terminal amidation, N-terminal acetylation, and methionine oxidation were simultaneously allowed as variable modifications. Peptide tolerance was set to 2 ppm, and MSMS tolerance was 25 millimass units. The significance threshold was the Mascot default setting of 5%. Each MSMS spectrum was checked manually to confirm or contradict the Mascot assignment. The false discovery rate for the identity threshold was in all cases 0% as estimated by using the Mascot decoy database function. The signals corresponding to intact calcitonin (CT), calcitonin gene-related peptide (CGRP), and somatostatin (SST) (with 1-ppm mass tolerance) underwent a mass shift of 116.01 Da after reductive alkylation with iodoacetamide and were sequenced as such by MSMS in a separate LC-MSMS analysis. Table I lists peptides that were identified with a score above the Mascot homology threshold. In the supplemental table, peptides with a score above the identity threshold (corresponding to an expectation value below 0.05) are listed and were considered identified.

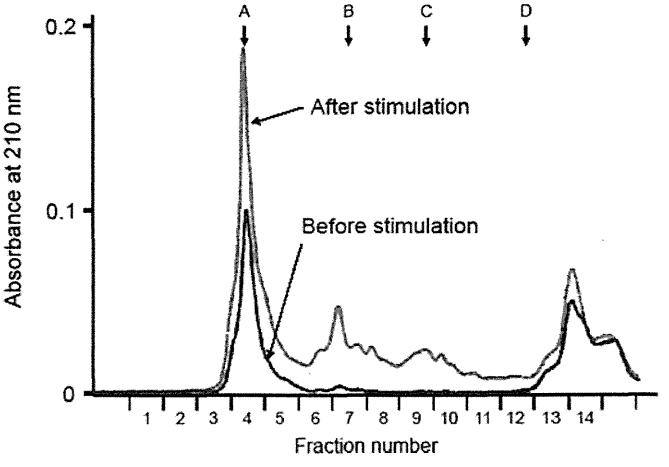

FIG. 1. **Gel filtration profiles of culture supernatant extracts from TT cells before (*black* trace) and after stimulation (*gray* trace).** *Arrows* indicate molecular mass markers: *A*, 66,500 Da; *B*, 4,271 Da; *C*, 1,673 Da; *D*, 556 Da.

RESULTS

Comprehensive Analysis of Peptides Released on Exocytosis—As a model system, we used the human medullary thyroid carcinoma cell line TT that stores peptide hormones including CT and CGRP in secretory granules (14, 15). A combination of forskolin and carbachol was used to induce exocytosis. Media from cells incubated for 2 min before and after stimulation were separately harvested and solid phase-extracted for peptide analysis. Total peptide amounts were assessed by gel filtration HPLC in which 1000–10,000-Da molecules are eluted in fractions labeled 7–10 (Fig. 1). This exocytotic stimulus elicited a 5.5-fold increase in secreted peptide amounts as assessed by the absorbance at 210 nm.

The solid phase-extracted samples were directly analyzed by LC-MSMS without gel filtration. We first examined a basal level secretion of peptides. In the medium conditioned by TT cells for 2 min, 36 peptides were identified from 13 precursors of which 30 peptides arose from nine secretory proteins including four peptide hormone precursors (CT/CGRP, gastrin-releasing peptide (GRP), and SST), four granin-like proteins (chromogranin A (CgA), chromogranin B (CgB), secretogranin III (SgIII), and VGF), and the processing enzyme PC2 (Fig. 2 and Table I). Because this cell line is known as a hyperproducer of CT and CGRP (14, 16), we tried to locate signals with mass values (within a mass tolerance of 2 ppm from a theoretical value) corresponding to bioactive CT (3415.58 Da) and CGRP (3786.96 Da). Signals from CGRP were observed, but no signals for CT were detected in the base peak chromatogram (Fig. 3).

In contrast, this stimulation facilitated identification of larger numbers of peptides, which resulted in 152 peptides being identified from 18 precursors (Fig. 2 and Table I). The six additional precursors all belonged to secretory proteins known to be processed in the RSP, including three peptide hormone precursors (pituitary adenylate cyclase-activating