Lung Cancer (2005) 49, 217–223

ELSEVIER

LUNG CANCER

IASLC

www.elsevier.com/locate/lungcan

Clinical responses of large cell neuroendocrine carcinoma of the lung to cisplatin-based chemotherapy

Shigeo Yamazaki[a,d], Ikuo Sekine[a,*], Yoshihiro Matsuno[b], Hidefumi Takei[c], Noboru Yamamoto[a], Hideo Kunitoh[a], Yuichiro Ohe[a], Tomohide Tamura[a], Tetsuro Kodama[a], Hisao Asamura[c], Ryosuke Tsuchiya[c], Nagahiro Saijo[a]

[a] *Division of Thoracic Oncology and Internal Medicine, National Cancer Center Hospital, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan*
[b] *Division of Clinical Laboratory, National Cancer Center Hospital, Tokyo*
[c] *Division of Thoracic Surgery, National Cancer Center Hospital, Tokyo*
[d] *Department of Surgery, Keiyu-kai Sapporo Hospital, Sapporo, Japan*

Received 30 September 2004; received in revised form 3 January 2005; accepted 3 January 2005

KEYWORDS
Neuroendocrine carcinoma;
Lung cancer;
Chemotherapy;
Cisplatin

Summary
Background: The efficacy of chemotherapy in patients with large cell neuroendocrine carcinoma of the lung (LCNEC) remains unclear.
Methods: Patients with LCNEC who received cisplatin-based chemotherapy were identified by reviewing 567 autopsied and 2790 surgically resected lung cancer patients. The clinical characteristics and objective responses to chemotherapy in these patients were analyzed.
Results: Overall, 20 cases of LCNEC were identified, including stage IIIA ($n=3$), stage IIIB ($n=6$), stage IV ($n=6$) and postoperative recurrence ($n=5$) cases. Six patients had received prior chemotherapy, and 14 were chemo-naive patients. The patients had received a combination of cisplatin and etoposide ($n=9$), cisplatin, vindesine and mitomycin ($n=6$), cisplatin and vindesine ($n=4$), or cisplatin alone ($n=1$). One patient showed complete response and nine showed partial response, yielding an objective response rate of 50%. The response rate did not differ between patients with the initial diagnosis of SCLC and those with the initial diagnosis of NSCLC, however, the response rate in chemo-naive patients (64%) was significantly different from that in previously treated patients (17%).
Conclusions: Our results suggest that the response rate of LCNEC to cisplatin-based chemotherapy was comparable to that of SCLC.

* Corresponding author. Tel.: +81 3 3542 2511; fax: +81 3 3542 3815.
E-mail address: isekine@ncc.go.jp (I. Sekine).

0169-5002/$ – see front matter
doi:10.1016/j.lungcan.2005.01.008

1. Introduction

Pulmonary neuroendocrine tumors include a spectrum of four clinicopathological entities classified on the basis of the morphological and biological features: typical carcionoid and atypical carcinoid, which are tumors of low to intermediate grade malignancy, and large cell neuroendocrine carcinoma (LCNEC) and small cell carcinoma (SCLC), which are high-grade malignant tumors. Travis et al. proposed the term LCNEC in 1991 [1], for classifying a type of poorly differentiated high-grade carcinoma characterized by a neuroendocrine appearance under light microscopy. LCNEC exhibits more prominent cellular pleomorphism and higher mitotic activity than the atypical carcinoid (AC), and is distinguished from SCLC by the tumor cell size and chromatin morphology. Although several different terminologies and classifications have been proposed previously, and even the present classification of pulmonary neuroendocrine tumors lacks uniform definition criteria, this class of tumors could become widely accepted and included in the updated histological classification of the World Health Organization [2].

The clinical features of LCNEC have not yet been completely clarified. The prognosis of patients with surgically resected LCNEC is reported to be intermediate between that of AC and SCLC [3–5], and the same as that of resected NSCLC, except that stage I LCNEC has a poorer prognosis than stage I non-small cell lung cancer (NSCLC) [6]. To the best of our knowledge, however, there are no studies that have examined the role of chemotherapy for LCNEC and the prognosis of patients with unresectable LCNEC, even though several reports have been published on the association between response to chemotherapy and the neuroendocrine differentiation of NSCLC [7–9]. The appropriate treatment of unresectable LCNEC, therefore, remains unclear. In the present study, we attempted to investigate the effectiveness of chemotherapy with cisplatin-based regimens for LCNEC in patients with unresectable and recurrent LCNEC.

2. Materials and methods

Eighty-seven of 2790 patients with primary lung cancer who underwent tumor resection from 1982 to 1999 at the National Cancer Center Hospital were found to have tumors with the histological characteristics of LCNEC [6]. Of these, five had received cisplatin-based chemotherapy at the time of recurrence, and were enrolled as subjects of this study. In addition, 303 of 567 patients who were autopsied from 1983 to 1997 at the National Cancer Center Hospital who had the following histological diagnoses were first selected: SCLC ($n=112$), poorly differentiated adenocarcinoma ($n=99$), large cell carcinoma ($n=58$), poorly differentiated squamous cell carcinoma ($n=29$), poorly differentiated adenosquamous carcinoma ($n=2$), LCNEC ($n=2$), and carcinoid ($n=1$). Of these, 161 had received cisplatin-based chemotherapy were selected for a pathological review. Finally, specimens from 17 of these161 cases were found to have histological characteristics consistent with the diagnosis of LCNEC, and were selected as subjects of this study. We focused on cisplatin, because since the 1980s, cisplatin has been the only anticancer agent with proven efficacy against both SCLC and NSCLC [10,11]; we, therefore, considered that the effectiveness of chemotherapy for LCNEC could be reasonably evaluated if cisplatin were included in the regimen. Cases which had received adjuvant chemotherapy without evaluable lesions were excluded from the analysis.

All the available paraffin-embedded tissue sections stained with hematoxylin–eosin were reviewed. We classified LCNEC according to the histopathological criteria in the WHO classification [2]. Immunohistochemical analysis was performed to confirm the neuroendocrine features of the tumors. For this purpose, formalin-fixed paraffin sections were stained for a panel of neuroendocrine markers, including chromogranin A (CGA), synaptophysin (SYN), and neural cell adhesion molecule (NCAM), using standard methods. The intensity of immunostaining for these markers was scored as follows: +, when the proportion of stained tumor cells was >50%; ±, when 10–50% of tumor cells were stained; and −, when <10% of tumor cells were stained, as previously described [6]. One case included in this study had the typical histological features of LCNEC, but no neuroendocrine features as determined by the immunohistochemical analysis. For specimens obtained after treatment, we routinely confirmed that the histopathological and morphological features showed no changes due to treatment as compared with the pretreatment biopsy or cytologic specimens. Such cases for which no pretreatment samples were available were excluded from the study; since it has been reported that histological changes may occur after treatment in SCLC [12], we were concerned that misdiagnosis might occur if the same were also true for LCNEC.

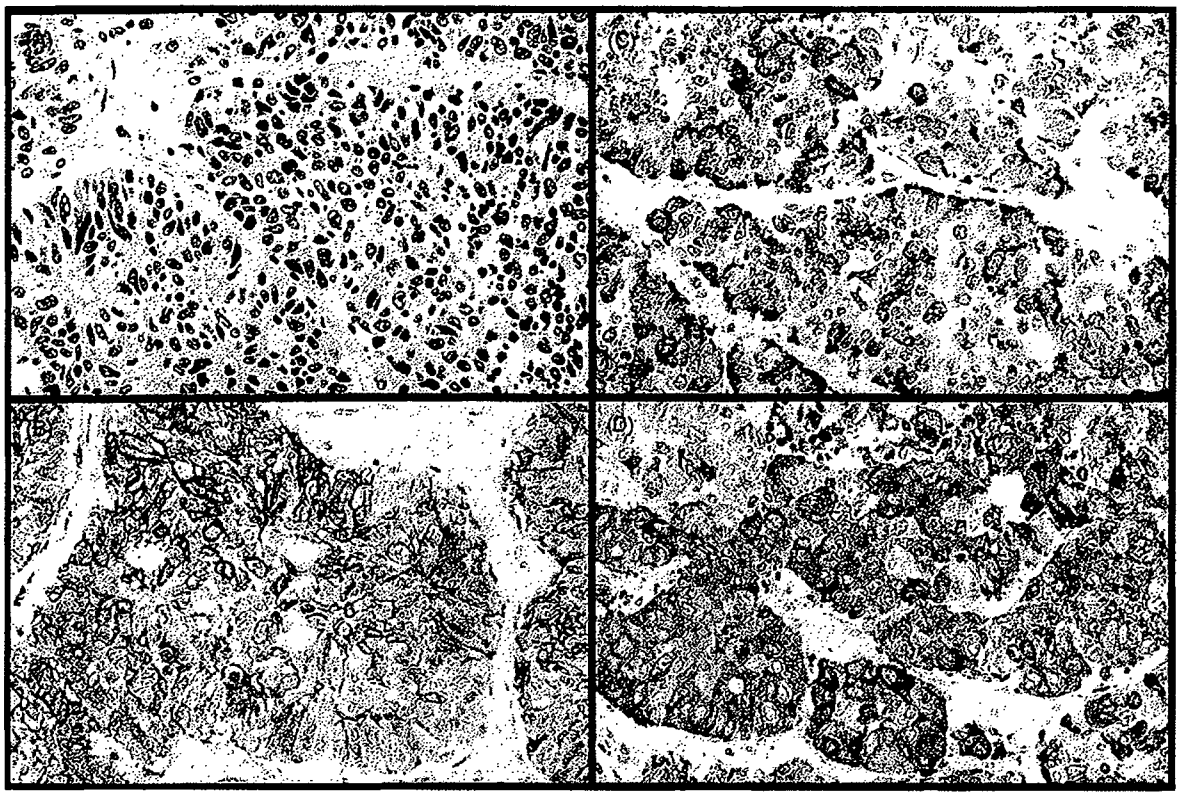

Fig. 1 Case no. 2, 57-year-old man. (A) The tumor cells which are large-sized, polygonal in shape and have a low nuclear-cytoplasmic ratio, are arranged in organoid nests and trabeculae (H&E stain, ×200). Positive staining for neural cell adhesion molecule (B), chromogranin A (C), and synaptophysin (D) (immunostain, ×400).

Clinical information about the cases was obtained from the medical records. The clinical disease staging was reassessed according to the latest International Union Against Cancer (UICC) staging criteria [13]. The response to chemotherapy and overall survival rate were assessed retrospectively. The objective tumor response was evaluated according to the WHO criteria published in 1979 (WHO, 1979) [14]. The survival time was measured from the date of start of chemotherapy with a cisplatin-containing regimen. Survival curves were drawn using the Kaplan–Meier method [15]. Drug toxicity could not be assessed as the study was a retrospective one and records were often incomplete.

3. Results

Overall, 22 cases were recognized as having tumors with histological characteristics consistent with LCNEC among the autopsied and surgically resected cases of primary lung cancer that had received cisplatin-based chemotherapy and had evaluable lesions; of these 17 were autopsied cases and five were surgically resected cases. Two of the autopsied cases were excluded, because no pretreatment pathological or cytological samples were available. The typical microscopic appearance of the tumor specimens is shown in Fig. 1A. The specimen sources for the prechemotherapy-diagnosis included surgically resected specimens ($n=5$), biopsy specimens ($n=9$), and cytology specimens ($n=6$). The histological and cytological findings in the specimens obtained before chemotherapy were consistent with those in the specimens obtained after chemotherapy. We therefore finally enrolled 20 cases in this study. The initial pathologic diagnoses in these patients were as follows: small cell carcinoma ($n=10$), poorly differentiated adenocarcinoma ($n=6$), large cell carcinoma ($n=2$), undifferentiated carcinoma ($n=1$), and poorly differentiated carcinoma ($n=1$) (Table 1). None of the cases had been labeled as LCNEC at the time of initial diagnosis, probably because the concept of LCNEC

Table 1 Patient characteristics

Characteristics	N	%
No. of patients	20	
Sex		
Male	18	90
Female	2	10
Age, median (range)	58 (37–74)	
Smoking history		
Yes	19	95
No	1	5
Performance status		
1–2	19	95
>2	1	5
Initial pathological diagnosis		
Small cell carcinoma	10	50
Adenocarcinoma	6	30
Large cell carcinoma	2	10
Others	2	10
Clinical stage at the start of chemotherapy		
IIIA	3	15
IIIB	6	30
IV	6	30
Postoperative recurrence	5	25
Prior treatment		
None	10	50
Surgery	4	20
Radiotherapy	2	10
Chemotherapy without cisplatin	6	30

Table 2 Staining for neuroendocrine markers in 20 LCNECs

Case	NCAM	CGA	SYN
1	+	+	+
2	+	+	+
3	+	+	+
4	±	+	+
5	+	+	+
6	+	+	+
7	−	+	−
8	+	−	−
9	−	−	−
10	−	+	±
11	+	−	+
12	+	+	+
13	+	+	+
14	+	−	±
15	+	+	+
16	+	−	NA
17	+	−	+
18	+	−	NA
19	+	−	+
20	−	+	+

NCAM, neural cell adhesion molecule; CGA, chromogranin A; SYN, synaptophysin; NA, not assessed.

was not completely accepted at our hospital at that time.

The results of the immunohistochemical staining are shown in Table 2, and a typical case showing positive staining is shown in Fig. 1B and D. Of the 20 LCNECs, 19 expressed at least one of the three general neuroendocrine markers, namely CGA, SYN, and NCAM. Sixteen of the 20 LCNECs exhibited positive staining for NCAM, while one showed equivocal staining. Twelve of the 20 LCNECs showed positive staining for CGA. Thirteen LCNECs showed positive staining for SYN and three showed equivocal staining. Only one case was negative for all the three general neuroendocrine markers, however, this case exhibited the typical histological features of LCNEC on light microscopy.

The clinical characteristics of the patients are summarized in Table 1. The extremely high predominance of men and smokers in this study was comparable to the demographic features of our LCNEC patients treated by surgical resection [6]. Previous chemotherapy was given in six patients: nedaplatin in one and cyclophosphamide-based regimen in five patients. The chemotherapy regimens used were as follows: cisplatin (80 mg/m^2, day 1) and etoposide (100 mg/m^2, days 1–3) (n = 9), cisplatin (80 mg/m^2, day 1), vindesine (3 mg/m^2, days 1 and 8) and mitomycin (8 mg/m^2, day 1) (n = 6), cisplatin (80 mg/m^2, day 1) and vindesine (3 mg/m^2, days 1 and 8) (n = 4), or cisplatin (100 mg/m^2, day 1) alone (n = 1). The median (range) number chemotherapy cycles were 2 (1–6). Of the 20 patients, one showed CR and nine showed PR, yielding an overall response rate of 50% (95% confidence interval, 27.2–72.8%). One CR and four PRs were observed among the cases treated with cisplatin and etoposide, two PRs were found among those treated with cisplatin, vindesine and mitomycin, and three PRs were found among those treated with cisplatin and vindesine. Seven patients showed NC, and three showed progressive disease. While the response rate did not differ between patients with an initial diagnosis of SCLC and those patients with an initial diagnosis of NSCLC, previous chemotherapy affected the response to cisplatin: the response rate in chemo-naive patients was 64%, whereas that in previously treated patients was 17%. The median progression-free survival in the 20 patients was 103 days, median survival was 239 days, 1-year survival rate was 35%, and 2-year survival rate was 15%.

4. Discussion

In this extensive review of over 3000 lung cancer patients, we found considerable difficulty in evaluating the response of LCNEC to systemic chemotherapy. The pathological diagnosis of LCNEC was established in 87 (3.1%) of 2790 patients treated by surgical resection. This low incidence of LCNEC in surgically treated lung cancer patients is comparable to that in other previously published reports: 2.4% (50/2070), 2.9% (22/766), and 3.6% (53/1530) [16–18]. Of the 87 patients, only five who had received cisplatin-based chemotherapy for recurrent tumor that was evaluable for the response. While LCNEC is difficult to diagnose prior to the start of treatment on the basis of the findings in biopsy or cytological specimens, the architectural neuroendocrine features may, more or less, be reflected in these small samples [19,20]. We, therefore, conducted a review of 567 autopsy cases of lung cancer, and identified 15 cases of LCNEC who had received cisplatin-based chemotherapy. We obtained a response rate to cisplatin-based chemotherapy of 50% in these 20 patients with LCNEC, however, the clinical characteristics of patients with medically treatable advanced LCNEC would still remain to be clarified, because autopsy is conducted only in highly selective cases.

Travis et al. suggested that immunohistochemical or electron-microscopic evidence of neuroendocrine features were important to diagnose LCNEC [1]. We assessed the neuroendocrine marker expression by immunohistochemical staining for CGA, SYN, and NCAM. Our cases included one that was negative for all the three neuroendocrine markers examined, but showed the typical histological features of LCNEC, which could be attributable to technical staining problems. Immunohistochemical staining for neuroendocrine tumors is generally recognized as only a supplementary diagnostic tool. In addition, the post-surgical survival rate did not differ between histologically diagnosed cases of LCNEC with neuroendocrine differentiation in marker expression as assessed by immunohistochemical staining and large cell carcinoma with neuroendocrine morphology where the neuroendocrine markers were negative (data not shown). Thus, we decided to include the case with negative staining as LCNEC on the basis of its typical neuroendocrine morphology.

To the best of our knowledge, only one study on the efficacy of chemotherapy in patients with LCNEC has been reported previously. In the study, 13 patients with LCNEC received chemotherapy when relapse was noted after surgical resection, and two (20%) of 10 evaluable patients showed an objective response. The evaluable lesion in these patients, however, was the brain in seven, liver in two, and bone in one patient [21]. Thus, the relatively low response rate in the report may be due to the site of the evaluable lesion. In addition, reports on the correlation between response to chemotherapy and neuroendocrine differentiation of NSCLC may be helpful. Graziano et al. reported that the proportion of NSCLC positive for neuroendocrine markers was higher in responders than in non-responders among 52 NSCLC patients treated by chemotherapy, and that the result suggested a correlation between positivity for neuroendocrine marker expression and the likelihood of response to chemotherapy [7]. On the other hand, others have reported the absence of any correlation between the presence of neuroendocrine differentiation and the response to chemotherapy [8,9]. The neuroendocrine differentiation in NSCLCs in the aforementioned studies was confirmed only by immunohistochemical staining and not on the basis of the morphological definition of LCNEC. Therefore, these groups might have potentially included heterogeneous subtypes of lung carcinoma, such as adenocarcinoma or squamous cell carcinoma, with components of neuroendocrine differentiation. The conflicting conclusions of these studies may, therefore, reflect differences in the biological characteristics of the tumors included in the analysis. Since the definition of LCNEC is based on morphological criteria as well as positivity for neuroendocrine marker expression, LCNEC is may be considered to be a clinically homogeneous group. Therefore, our study of LCNEC may endorse the former reports about the relationship between neuroendocrine differentiation and the sensitivity to chemotherapy.

Objective response to chemotherapy can be observed in only 15–30% of NSCLCs, even when they are treated with regimens containing cisplatin [10]. In SCLC, however, effective combination regimens yield objective response rates in the range of 80–90% [11]. Our study showed an overall response rate of LCNEC of 50% to cisplatin-based chemotherapy, and a response rate of 64% in chemo-naive patients, which seemed to be higher than the response rate of NSCLC to chemotherapy. Considered together, these results suggest that the chemosensitivity of LCNEC is intermediate between that of NSCLC and SCLC, although we were unable to obtain firm evidence from this retrospective study, which included only a small cohort of patients.

Since LCNEC is a relatively rare subtype of lung cancer, a prospective study is difficult to perform, and may only be possible as a multicenter study.

For this purpose, it is an urgent task to establish diagnostic criteria for LCNEC based on examination of biopsy or cytologic specimens. Although the histological definition of LCNEC in surgically resected specimens proposed by Travis et al. is commonly accepted, its diagnostic reproducibility is not satisfactory [22]. It is also difficult to apply the definition to biopsy specimens, in which artifacts can easily be produced and detailed examination may be difficult due to insufficient specimen size. Thus, definitive diagnostic criteria also applicable to biopsy and cytologic specimens are required.

Our study did not include any cases labeled as LCNEC at the time of initial diagnosis. One half of the cases was originally diagnosed as SCLC and the other half as NSCLC, including poorly differentiated adenocarcinoma and large cell carcinoma. This was attributed to the fact that the concept of LCNEC was not clearly defined prior to its being proposed by Travis et al. [1]. Thus, it is possible that patients with LCNEC were included in earlier clinical trials for NSCLC or SCLC. If LCNEC shares the poor prognosis of NSCLC, the reported results of chemotherapy for NSCLC may have been worse in studies in which cases of LCNEC were included. Similarly, the results of clinical studies of SCLC to study their objective response to chemotherapy may also have been worse because of the confounding effects of the inclusion of LCNECs among the cases.

In conclusion, our results suggest that the response rate of LCNEC to cisplatin-based chemotherapy was comparable to that of SCLC. However, because of the retrospective nature of this study and the small sample size, we could not arrive at any definitive conclusion; we, therefore, propose to conduct a prospective study in the future aimed at elucidating the efficacy of chemotherapy for LCNEC. To that end, firm diagnostic criteria for LCNEC need to be established, even when the diagnosis must be based only on examination of biopsy and cytology specimens.

Acknowledgment

We thank Ms. Yuko Yabe for kindly preparing this manuscript.

References

[1] Travis WD, Linnoila RI, Tsokos MG, Hitchcock CL, Cutler Jr GB, Nieman L, et al. Neuroendocrine tumors of the lung with proposed criteria for large-cell neuroendocrine carcinoma. An ultrastructural, immunohistochemical, and flow cytometric study of 35 cases. Am J Surg Pathol 1991;15:529–53.

[2] Travis W, Corrin B, Shimosato Y, Brambilla E. Histological typing of lung and pleural tumours. Berlin: Springer-Verlag; 1999.

[3] Rusch VW, Klimstra DS, Venkatraman ES. Molecular markers help characterize neuroendocrine lung tumors. Ann Thorac Surg 1996;62:798–809 [discussions 09–10].

[4] Dresler CM, Ritter JH, Patterson GA, Ross E, Bailey MS, Wick MR. Clinical-pathologic analysis of 40 patients with large cell neuroendocrine carcinoma of the lung. Ann Thorac Surg 1997;63:180–5.

[5] Travis WD, Rush W, Flieder DB, Falk R, Fleming MV, Gal AA, et al. Survival analysis of 200 pulmonary neuroendocrine tumors with clarification of criteria for atypical carcinoid and its separation from typical carcinoid. Am J Surg Pathol 1998;22:934–44.

[6] Takei H, Asamura H, Maeshima A, Suzuki K, Kondo H, Niki T, et al. Large cell neuroendocrine carcinoma of the lung: a clinicopathologic study of eighty-seven cases. J Thorac Cardiovasc Surg 2002;124:285–92.

[7] Graziano SL, Mazid R, Newman N, Tatum A, Oler A, Mortimer JA, et al. The use of neuroendocrine immunoperoxidase markers to predict chemotherapy response in patients with non-small-cell lung cancer. J Clin Oncol 1989;7:1398–406.

[8] Schleusener JT, Tazelaar HD, Jung SH, Cha SS, Cera PJ, Myers JL, et al. Neuroendocrine differentiation is an independent prognostic factor in chemotherapy-treated nonsmall cell lung carcinoma. Cancer 1996;77:1284–91.

[9] Carles J, Rosell R, Ariza A, Pellicer I, Sanchez JJ, Fernandez-Vasalo G, et al. Neuroendocrine differentiation as a prognostic factor in non-small cell lung cancer. Lung Cancer 1993;10:209–19.

[10] Ginsberg RJ, Vokes EE, Rosenzweig K. Non-small cell lung cancer. In: DeVita VT, Hellman S, Rosenberg SA, editors. Cancer: principles & practice of oncology. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 925–83.

[11] Murren J, Glatstein E, Pass HI. Small cell lung cancer. In: DeVita VT, Hellman S, Rosenberg SA, editors. Cancer: principles & practice of oncology. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 983–1018.

[12] Fushimi H, Kikui M, Morino H, Yamamoto S, Tateishi R, Wada A, et al. Histologic changes in small cell lung carcinoma after treatment. Cancer 1996;77:278–83.

[13] Sobin LH, Wittenkind C. TNM classification of malignant tumors. New York: Wiley-Liss; 1997.

[14] Handbook for reporting results of cancer treatment. Geneva: World Health Organization (WHO); 1979.

[15] Kaplan ELMP. Non-parametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457–81.

[16] Jiang SX, Kameya T, Shoji M, Dobashi Y, Shinada J, Yoshimura H. Large cell neuroendocrine carcinoma of the lung: a histologic and immunohistochemical study of 22 cases. Am J Surg Pathol 1998;22:526–37.

[17] Iyoda A, Hiroshima K, Toyozaki T, Haga Y, Fujisawa T, Ohwada H. Clinical characterization of pulmonary large cell neuroendocrine carcinoma and large cell carcinoma with neuroendocrine morphology. Cancer 2001;91:1992–2000.

[18] Paci M, Cavazza A, Annessi V, Putrino I, Ferrari G, De Franco S, et al. Large cell neuroendocrine carcinoma of the lung: a 10-year clinicopathologic retrospective study. Ann Thorac Surg 2004;77:1163–7.

[19] Yang YJ, Steele CT, Ou XL, Snyder KP, Kohman LJ. Diagnosis of high-grade pulmonary neuroendocrine carcinoma by fine-needle aspiration biopsy: nonsmall-cell or small-cell type? Diagn Cytopathol 2001;25:292–300.

[20] Wiatrowska BA, Krol J, Zakowski MF. Large-cell neuroendocrine carcinoma of the lung: proposed criteria for cytologic diagnosis. Diagn Cytopathol 2001;24:58–64.
[21] Mazieres J, Daste G, Molinier L, Berjaud J, Dahan M, Delsol M, et al. Large cell neuroendocrine carcinoma of the lung: pathological study and clinical outcome of 18 resected cases. Lung Cancer 2002;37:287–92.
[22] Travis WD, Gal AA, Colby TV, Klimstra DS, Falk R, Koss MN. Reproducibility of neuroendocrine lung tumor classification. Hum Pathol 1998;29:272–9.

REGULAR ARTICLE

Clinical-scale high-throughput human plasma proteome analysis: Lung adenocarcinoma

Kiyonaga Fujii[1], Tomoyo Nakano[1], Mitsuhiro Kanazawa[3], Shingo Akimoto[3], Takashi Hirano[2], Harubumi Kato[2] and Toshihide Nishimura[1, 3]

[1] Clinical Proteome Center
[2] Department of Surgery, Tokyo Medical University, Shinjuku, Tokyo, Japan
[3] Medical ProteoScope Co., Ltd., Shinjuku, Tokyo, Japan

Clinical proteomics requires the stable and reproducible analysis of a large number of human samples. We report a high-throughput comprehensive protein profiling system comprising a fully automated, on-line, two-dimensional microflow liquid chromatography/tandem mass spectrometry (2-D μLC-MS/MS) system for use in clinical proteomics. A linear ion-trap mass spectrometer (ITMS) also known as a 2-D ITMS instrument, which is characterized by high scan speed, was incorporated into the μLC-MS/MS system in order to obtain highly improved sensitivity and resolution in MS/MS acquisition. This system was used to evaluate bovine serum albumin and human 26S proteasome. Application of these high-throughput μLC conditions and the 2-D ITMS resulted in a 10-fold increase in sensitivity in protein identification. Additionally, peptide fragments from the 26S proteasome were identified three-fold more efficiently than by the conventional 3-D ITMS instrument. In this study, the 2-D μLC-MS/MS system that uses linear 2-D ITMS has been applied for the plasma proteome analysis of a few samples from healthy individuals and lung adenocarcinoma patients. Using the 2-D and 1-D μLC-MS/MS analyses, approximately 250 and 100 different proteins were detected, respectively, in each HSA- and IgG-depleted sample, which corresponds to only 0.4 μL of blood plasma. Automatic operation enabled the completion of a single run of the entire 1-D and 2-D μLC-MS/MS analyses within 11 h. Investigation of the data extracted from the protein identification datasets of both healthy and adenocarcinoma groups revealed that several of the group-specific proteins could be candidate protein disease markers expressed in the human blood plasma. Consequently, it was demonstrated that this high-throughput μLC-MS/MS protein profiling system would be practically applicable to the discovery of protein disease markers, which is the primary objective in clinical plasma proteome projects.

Received: May 14, 2004
Revised: October 13, 2004
Accepted: November 15, 2004

Keywords:
Clinical Proteomics / Human plasma / Linear ion-trap mass spectrometry / Lung adenocarcinoma / Plasma proteome

Correspondence: Professor Toshihide Nishimura, Professor of Clinical Proteome Center, Tokyo Medical University, 2-6-1, Nishi-shinjuku, Shinjuku-ku, Tokyo 163-0217, Japan
E-mail: nisimura@tokyo-med.ac.jp
Fax: +1-81-3-5321-6624

Abbreviations: ABC, ammonium bicarbonate; **AID-HP**, albumin-and-IgG-depleted human plasma; **IAM**, iodoacetamide; **Ig**, immunoglobulin; **ITMS**, ion-trap mass spectrometry; **μLC**, microflow liquid chromatography; **NSI**, nanoelectrospray ionization; **SCX**, strong cation exchange; **TCEP**, tris[2-carboxyethyl]phosphine; **TPX**, methylpentene polymer

1 Introduction

Human blood plasma is generally the most informative proteome from a medical viewpoint, because it is the primary clinical specimen and it also represents the largest and deepest version of human proteome present in any sample [1–3]. Almost all body cells communicate with the plasma either directly or through tissues/biological fluids, and many of these cells release at least a part of their contents into the plasma upon damage or death. A comprehensive, systematic characterization of the plasma proteome in the healthy and

diseased states will greatly facilitate the development of biomarkers for early disease detection, clinical diagnosis, and therapy of cancer and other diseases. However, broad characterization of the human plasma proteome may pose one of the greatest challenges. This is because it can contain low-level proteins, which are secreted by solid tissues, as well as other important proteins (tissue leakage proteins at pg/mL levels) in the presence of several relatively dominant, high-abundance proteins (particularly HSA at 35–50 mg/mL). The dynamic range of plasma protein concentrations minimally spans nine orders of magnitude. For clinical and diagnostic proteomics using human plasma, it is essential to develop a comprehensive system, which has a high resolution and a wide dynamic range, for large-scale proteome analysis.

Recently, multi-dimensional LC-MS/MS has been developed as a powerful tool, particularly for comprehensive identification of highly complex proteins. This method can achieve a resolving power that is equal to or higher than 2-DE [4–6]. Broad protein identification techniques can detect specific proteins present in low concentrations in a highly complex protein matrix. To characterize the human plasma proteome, Smith *et al.* have achieved a protein identification dynamic range of more than eight orders of magnitude using 2-D LC combined with conventional ion-trap MS/MS instrumentation [6]. This approach has resulted in the identification of >800 plasma proteins from 5 μL plasma without the depletion of highly abundant HSA and/or immunoglobulins (Ig). The multi-dimensional LC-MS/MS techniques reported thus far indicate the potential usefulness of broad protein identification with high resolution and wide dynamic range for cataloging the plasma contents. However, these approaches require further improvement in terms of both ease of use and industrial applicability to routine clinical use, because their application to clinical research requires stable and reproducible analyses of a large number of human samples.

The establishment of a simple, robust, and high-throughput protein profiling system as a global platform is extremely important from the viewpoint of clinical proteomics. This is because a large number of human tissue/biological fluid samples could then be quantitatively analyzed, in a routine and reproducible manner, for expressed proteins. Such a system would help discover any protein that is significantly associated with a specific disease status. We have constructed a technically well integrated and high-throughput LC-MS/MS system with RP microflow LC (μLC) and a conventional ion-trap MS/MS equipped with a nanoelectrospray ionization (NSI) interface to detect lung cancer biomarkers and to analyze apoptotic mechanisms [7, 8]. Additionally, the system has been combined with on- or off-line strong cation-exchange (SCX) chromatography to result in a multi-dimensional protein profiling system. This protein profiling system using off-line 2-D SCX/RP μLC-MS/MS was successfully applied to broad protein identification of human plasma proteins [9]. We have also established protein depletion, in-solution digestion, and data-integrating/mining systems with an automated operation for large-scale human plasma proteome analysis.

The dynamic range and sequence coverage that results from protein identification by LC-MS/MS analysis depends on both the quality of the separation(s) applied and the MS platform [6]. When resolution performance is not considered, the quality of the 2-D LC is substantially related to the number of fractionation steps for the first dimension chromatography and the analytical running time for the second dimension chromatography. Since the dynamic range of the MS platform is based on the performance of the instrument used, the number of the MS/MS acquisition in one run strictly depends on both the scan speed and the analytical time required by the LC-MS/MS analysis. Recently, linear ion-trap MS/MS (2-D ITMS) instruments with a higher scan speed and sensitivity than conventional 3-D ITMS instruments have emerged as new generation instruments. Therefore, we applied the new 2-D ITMS instrument to the fully automated on-line 2-D μLC-MS/MS system developed by us [10].

In this study, we evaluated the performance of the μLC-MS/MS system using the 2-D ITMS instrument for extending the sensitivity, dynamic range, and coverage for comprehensive protein identification. BSA and human 26S proteasome were used as the authentic protein sample and the protein complex sample, respectively. The system, together with the on-line 2-D μLC-MS/MS system, was then applied to proteome analysis of human plasma; HSA- and IgG-depleted samples were obtained from a few healthy individuals and lung adenocarcinoma cases.

2 Materials and methods

2.1 Materials

HPLC-grade ACN, formic acid, and TFA were purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Milli-Q grade water (Millipore, Bedford, MA, USA) was used. BSA, ammonium formate, ammonium bicarbonate (ABC), and iodoacetamide (IAM) were purchased from Sigma (St. Louis, MO, USA). Human 26S proteasome (PW9310) was obtained from Affiniti Research Products (Devon, UK). Tris[2-carboxyethyl]phosphine (TCEP) was purchased from Pierce (Rockford, IL, USA). Sequencing grade-modified trypsin was purchased from Promega (Madison, WI, USA).

2.2 Preparation of the digested BSA and human 26S proteasome samples

BSA (1 nmol) was diluted with 225 μL ABC (aq., 100 mM); then, 12.5 μL TCEP (10 mM) was added for reduction and the solution mixture was kept at 37°C for 45 min. Further, 12.5 μL IAM (50 mM) was added, and the solution mixture was alkylated in the dark at 24°C for 1 h. The resulting solution was digested with trypsin (trypsin:protein = 1:50, w/w),

and the resultant 250 μL solution was incubated in the dark at 37°C for 15 h. In-solution digestion of the 26S proteasome sample was carried out as follows: 26S proteasome sample (50 μg) was diluted with ABC (aq., 50 mM) containing 10% v/v ACN to a final volume of approximately 190 μL. For reduction, 2.5 μL TCEP (10 mM) was added, and the solution mixture was kept at 37°C for 45 min. Subsequently, 2.5 μL IAM (50 mM) was added, and the solution mixture was alkylated in the dark at 24°C for 1 h. For digestion, trypsin (2 μg) was added to 5 μL ABC (50 mM), and 200 μL of the resulting solution was incubated in the dark at 37°C for 18 h. All reactions were performed in methylpentene polymer (TPX) microtubes (Hitech Inc. Tokyo, Japan) using an Eppendorf thermomixer R (Brinkmann, Westbury, NY, USA) for 1.5 mL microtubes; the resulting solution was interval-mixed (10 s) at 850 rpm. The digested 26S proteasome sample (50 μL) was diluted with 125 μL of 2% v/v ACN (aq.) containing 0.005% v/v TFA (aq.) after adjusting the pH to approximately 3 with 50 μL of 1% v/v TFA (aq.); the samples (25 μL) were then injected into the μLC-MS/MS system described in this paper.

2.3 Sample preparation of the digested human plasma protein mixture

The human blood plasma samples treated with heparin were obtained from Tokyo Medical University (Tokyo, Japan) and acquired from three, healthy, anonymous, male donors (samples: H-N, H-I, and H-S) and two male donors who were diagnosed with adenocarcinoma on the basis of clinical and laboratory criteria (stage: IIIA, samples: AC88 and AC94), after obtaining their informed consent. HSA and IgG in the human plasma samples (500 μL) was removed by affinity adsorption chromatography using Bio-Rad's Affi-Gel Blue Gel and protein A column (Bio-Rad Hercules, CA, USA), respectively (details not shown). The final concentration of the resulting HSA- and IgG-depleted human plasma (AID-HP) samples were 4.1 (H-N), 8.6 (H-I), 6.6 (H-S), 5.8 (AC88), and 7.2 mg/mL (AC94) in ABC (25 mM). Subsequently, 100 μL of the AID-HP sample was diluted with 400 μL ABC (25 mM) containing 32% v/v ACN. For reduction, 25 μL TCEP (50 mM) was added and the solution mixture was kept at 37°C for 45 min. Subsequently, 25 μL IAM (250 mM) was added and the solution mixture was alkylated in the dark at 24°C for 1 h. For digestion, trypsin (5 μg) was added and the resulting solution (555 μL) was incubated in the dark at 37°C for 16 h. All these reactions were carried out using the Eppendorf thermomixer R for 1.5 mL TPX microtubes, and the mixing was carried out at 850 rpm with periods and intervals of 10 s each. For the 1-D μLC-MS/MS analysis, the digested AID-HP samples (20 μL) were diluted with 20 μL of 1% v/v TFA (aq.) and 160 μL with 2% v/v ACN (aq.) containing 0.1% v/v TFA, in a TPX auto sampler tube; 20 μL of the resultant samples was injected into the system. In order to prepare individual mixture samples of the healthy and adenocarcinoma groups for 1-D and 2-D μLC-MS/MS analyses, 33.3 μL of each of the three AID-HP samples from the healthy group and 50 μL of each of the two AID-HP samples from the adenocarcinoma group were mixed with 100 μL of 1% v/v TFA (aq.), respectively; subsequently, 4 μL of the resulting sample solutions was used in these analyses.

2.4 1-D RP and 2-D SCX/RP μLC-NSI-MS/MS analyses

The 1-D and 2-D LC-MS/MS system with RP-μLC comprised a Paradigm MS4 dual solvent delivery system (Michrom BioResources, Auburn, CA, USA) for HPLC, an HTS PAL auto sampler with two 10-port injector valves (CTC Analytics, Zwingen, Switzerland), Finnigan LCQ Deca XP plus 3-D ion-trap, and Finnigan LTQ linear ITMS (Thermo Electron, San Jose, CA, USA) equipped with NSI sources (AMR Inc., Tokyo, Japan). Sample injection for the 1-D and 2-D μLC-MS/MS analyses as well as SCX separation for 2-D analysis were automatically carried out using the HTS PAL auto sample injection system with no change in the configurations. The SCX separation was performed on an SCX microtrap cartridge (12 μm, 300 Å, 8 × 1.0 mm i.d., Michrom) by stepwise elution on the first injector valve. The solvent system containing 2% v/v ACN was composed of 0.005% v/v TFA (aq.) and 1 M ammonium formate (aq.) adjusted to pH 2.8 with TFA, and the elution solvents (25, 50, 100, 150, 200, and 500 mM) were prepared by mixing these. The effluent from all the SCX fractions was flowed serially into a peptide CapTrap cartridge (2.0 × 0.5 mm i.d., Michrom), present on the second injector valve, for concentration and desalting. After desalting with 0.1% v/v TFA (aq.) containing 2% v/v ACN, the sample was loaded onto a capillary RP column, MAGIC C_{18} (3 μm, 200 Å, 50 × 0.2 mm i.d., Michrom), for 2-D separation. Digested samples for the 1-D RP analysis were also injected directly into a peptide CapTrap cartridge for concentration and desalting and then applied to RP separation. Solutions of 2% and 90% v/v ACN (aq.) were used as the mobile phases A and B, respectively, and both contained 0.1% v/v formic acid. The gradient conditions in the chromatographic run were as follows: B 5% (0 min) → 65% (20 min) for the digested samples of BSA and 26S proteasome, and B 5% (0 min) → 40% (70 min) → 95% (80 min) for plasma samples. Effluent solvent at 1.0–1.2 μL/min from the HPLC was introduced into the mass spectrometer by the NSI interface *via* an injector valve with a CapTrap cartridge and the RP column. The NSI needle (FortisTip, OmniSeparo-TJ, Hyogo, Japan), which was connected directly to the RP column outlet, was used as the NSI interface and the voltage was 1.8 kV, while the capillary was heated to 200°C [11]. No sheath or auxiliary gas was used. Further, the mass spectrometer was operated in a data-dependent acquisition mode in which MS acquisition with a mass range of m/z 450–2000 was automatically switched to MS/MS acquisition under the automated control of the Xcalibur software. The most intense ion of the full MS scan was selected as the parent ion and it was subjected to MS/MS scan with an isolation width of m/z 2.0; the activation amplitude parameter

was set at 30%. For the human plasma samples, the full MS scan was acquired followed by two successive MS/MS scans of the two most intense precursor ions detected in the full MS scan. The trapping time was 100 ms under the auto gain control mode. Data was acquired using the dynamic mass-exclusion windows that had an exclusion of 3.0-min duration and exclusion mass widths of −0.5 and +1.5 Da.

2.5 Database searches

All MS/MS data were investigated using the Mascot search engine (Matrix Science, London, UK) [12] against the Swiss-Prot database. The data acquired for BSA digests were investigated against other mammalian subsets of the sequences. The MS/MS data of the human 26S proteasome and plasma samples were investigated against the *Homo sapien* subsets of the sequences. The database searches allowed for fixed modification on the cysteine residue (carbamidomethylation, +57 Da), variable modification on the methionine residue (oxidation, +16 Da), peptide mass tolerance at ±2.0 Da, and fragment mass tolerance at ±0.8 Da.

3 Results and discussion

3.1 Evaluation of the μLC-MS/MS analysis using the linear 2-D ITMS instrument

We developed the μLC-MS/MS system with RP separation (1-D RP), which corresponds to the second dimension separation for the on-line and off-line 2-D μLC-MS/MS system [9, 10]. This system comprises a microflow LC system with a variable splitter, a versatile auto-sampler equipped with an injector valve, and a LCQ 3-D (ITMS) with an NSI stage. A flow rate of 1.0–1.2 μL/min *via* the injector valve and the RP column (0.2 mm i.d.) has been adopted as a convenient and efficient condition for routine proteome analysis with high sensitivity and reproducibility. The detection limit for identification of proteins in protein digests was approximately a few fmol. In order to evaluate the sensitivity of the new linear 2-D ITMS instrument (LTQ) for protein identification, we connected our RP μLC system to the LTQ instead of the conventional 3-D ITMS (LCQ) instrument. BSA digests (5–500 fmol) were applied to μLC-MS/MS analysis using the LTQ and LCQ instruments under identical conditions except those used for the mass spectrometer. The base-peak chromatogram for BSA digests (500 fmol) is shown in Fig. 1a. A comparison of the coverage, in terms of protein identification, between the LCQ and LTQ instruments revealed that 25% coverage of the BSA sequence was acquired from 5 fmol of the digests using the LTQ instrument, as shown in Fig. 2. Since the same coverage was obtained from 50 fmol of the digests using the LCQ instrument, the results indicated that the protein identification improved markedly as the sensitivity increased 10-fold using the LTQ instrument.

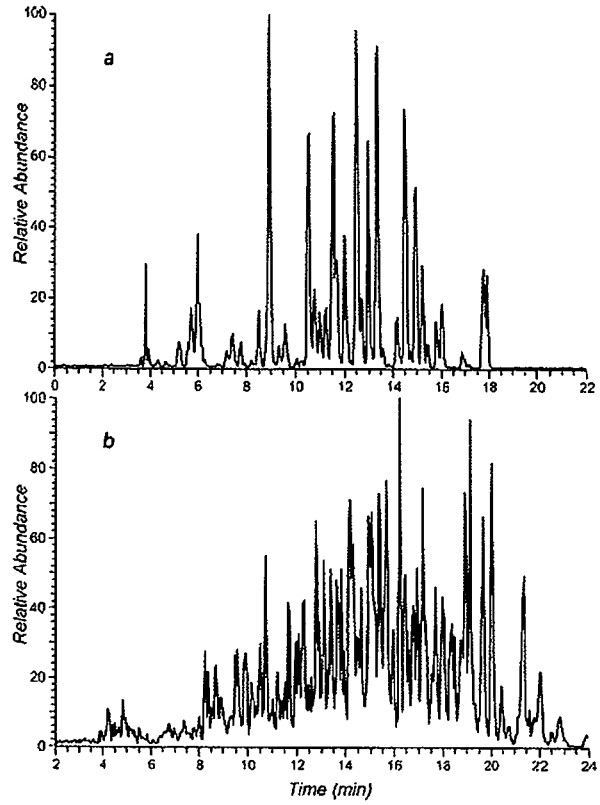

Figure 1. Base-peak chromatograms of the digested BSA (a) and 26S proteasome (b) using 1-D RP μLC-MS/MS analysis.

A dramatic improvement is achieved in the LTQ instrument in terms of the scan speed, which is higher than that of conventional 3-D ITMS instruments. Figures 3a and 3b show the expanded mass chromatograms with stick plotting at *m/z* values 710.0–711.5 obtained by μLC-MS/MS analysis of BSA digests using the LTQ and LCQ instruments, respectively. A stick in the peak represents a single full MS or MS/MS scan. The range denoted by the arrow in Fig. 3 shows that 62 events that carried out the acquisition of a full MS and an MS/MS spectra were achieved by the LTQ instrument in comparison with 12 events acquired by the LCQ instrument in 30 s. When conventional LCQ instruments are used, we usually apply three and two microscans for full MS (50 ms trapping time) and MS/MS (200 ms) accumulations, respectively, in order to obtain a better quality spectrum from a single scan. On the other hand, both spectra for peptide sequencing in protein identification were effectively acquired by one microscan of both full MS (50 ms) and MS/MS (100 ms) accumulations by LTQ instrument. As a result, using the LTQ instrument, it is possible to obtain an approximately five-fold higher number of MS/MS spectra in the same analytical run time.

In general, it is necessary to acquire a greater number of MS/MS spectra for the identification of a greater number of proteins using peptide sequencing. In studies where LC-MS/

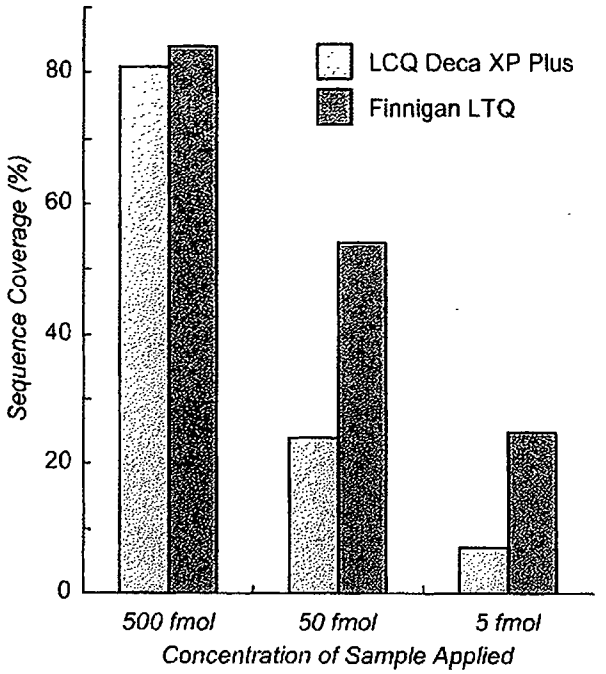

Figure 2. Comparison of sequence coverage of BSA digests (5–500 fmol) by μLC-MS/MS analysis using LCQ Deca XP plus 3-D ion-trap and Finnigan LTQ linear ion-trap MS instruments.

MS was used for comprehensive proteome analysis, several researchers investigated various methods to data-dependently amass MS/MS spectra for a single analysis. These methods include employing a longer analytical time, triple and more MS/MS acquisition against a single full MS spectrum, multiple analyses of the same sample using common conditions or fractionated mass range, etc. While obtaining the MS/MS spectra, the scan speed is a mechanical limitation and a data-dependent scan misses many of the peptide sequences for low abundance peaks that are behind large peaks. Therefore, it is necessary to choose the applications of these techniques for comprehensive proteome analysis of highly complex protein mixtures such as human plasma and whole cell lysates. The drawbacks of clinical proteomics for a large number of human samples are the inability to conduct multiple analyses of the same sample and the longer running time required by the comprehensive LC-MS/MS analysis. Consequently, the performance of the LTQ instrument with a higher scan speed is better than that of the conventional 3-D ITMS instrument, because it enables more informative high-throughput LC-MS/MS analysis for highly complex clinical samples. Therefore, to verify the applicability of LTQ instruments, the human 26S proteasome, which is a highly complex protein mixture consisting of 31 components, was subjected to analysis. As shown in Fig. 1b, equivalent amounts of the digested 26S proteasome sample were analyzed using the LCQ and LTQ instruments under identical μLC conditions. The data-dependent MS/MS

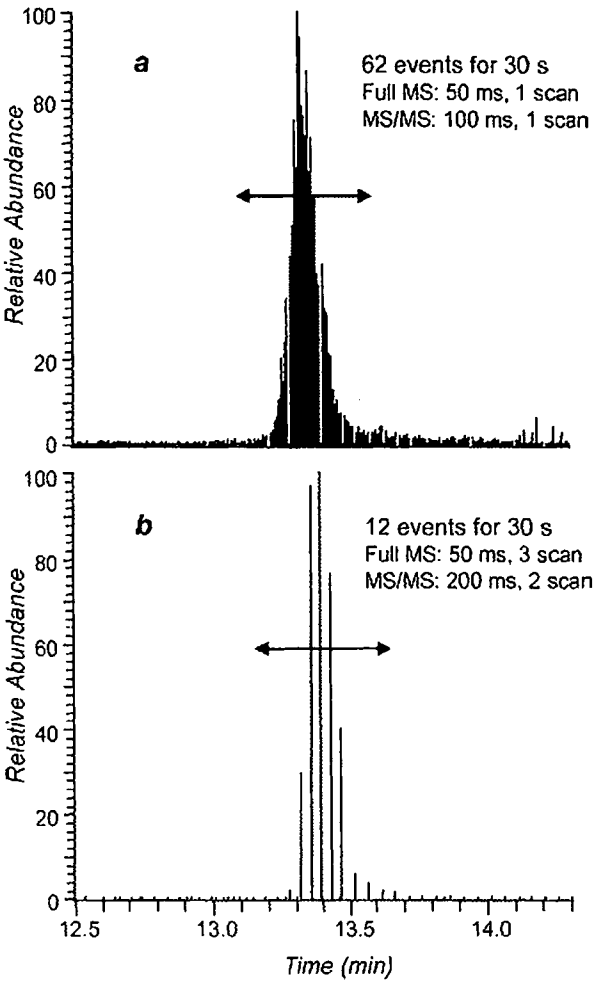

Figure 3. Mass chromatograms at *m/z* 710.0-711.5 by μLC-MS/MS analysis of BSA digests using Finnigan LTQ (a) and LCQ Deca XP plus (b) instruments. A stick in the peak is a single full MS and MS/MS scan of the mass spectrometer, and 62 and 12 events (acquisition of a full MS and a MS/MS in one event) were carried out by the LTQ and LCQ instruments, respectively.

acquisition, in which the full MS acquisition is followed by a single MS/MS scan of the most intense precursor ion obtained from the full MS scan, was applied with three microscan full MS (50 ms trapping time) and two microscan MS/MS (200 ms) accumulations for the LCQ, and one microscan of both full MS (50 ms) and MS/MS (100 ms) accumulations for the LTQ instruments. During the 20 min analysis, approximately 450 and 2200 MS/MS spectra were obtained from the 1-D RP μLC-MS/MS analyses using the LCQ and LTQ instruments, respectively. These data were evaluated using a Mascot database search against the Swiss-Prot database, and the search results obtained for the peptide MS/MS assignment were filtered based on the criterion defined as a Mascot peptide score more than 20 and ranked first, described in detail below. Table 1 (see also Supplemen-

Table 1. Protein identification results of human 26S proteasome.

Protein	LCQ Deca XP Plus			Finnigan LTQ		
	Score	Coverage	Peptide	Score	Coverage	Peptide
26S protease regulatory subunits						
PRS4	221	13	3	851	41	16
PRS6	142	9	3	757	48	16
PRS7	348	15	5	1346	55	23
PRS8	299	16	4	1304	58	20
PRSA	377	21	7	1093	53	19
PRSX	154	7	2	715	36	12
Proteasome subunit alpha types						
PSA1	494	36	8	658	48	11
PSA2	261	18	4	569	56	9
PSA3	145	13	3	584	43	11
PSA4	115	9	2	459	33	7
PSA5	158	17	3	483	47	8
PSA6	559	38	8	700	48	11
PSA7	423	35	7	772	57	12
Proteasome subunit beta types						
PSB1	275	34	5	596	53	10
PSB2	256	24	4	456	43	8
PSB3	235	24	3	379	35	5
PSB4	117	9	2	511	48	8
PSB5	528	42	7	670	58	10
PSB6	200	13	3	271	24	5
PSB7	232	14	4	381	33	7
PSB8	ND	ND	ND	180	18	4
26S proteasome non-ATPase regulatory subunits						
PSD1	227	6	3	1636	38	27
PSD2	380	10	6	1386	34	26
PSD3	690	24	10	1445	50	25
PSD4	74	5	1	358	22	6
PSD6	206	12	3	1001	46	18
PSD7	250	15	3	487	42	10
PSD8	ND	ND	ND	228	19	5
PSDB	571	22	8	1445	59	23
PSDC	359	15	5	1147	37	20
PSDD	183	10	4	966	52	17
		Total	130		Total	409

The protein identification data of 31 components consisted of human 26S proteasome including the number of the peptide fragments assigned (Peptide) and the sequence coverage according to these peptides (Coverage). The protein score is calculated by the addition of these peptide scores (Score) in comparison to 1-D μLC-MS/MS analysis using conventional 3-D ion-trap MS (LCQ Deca XP Plus) and new linear ion-trap MS instruments (Finnigan LTQ). ND, not detected.

tary Table A) shows protein identification results including the number of peptide fragments assigned, sequence coverage with these peptides, and the protein score calculated by addition of these peptide scores with respect to the 31 components of the 26S proteasome. In the case of the LTQ instrument, 409 peptide fragments were assigned as components of the 26S proteasome, and this number was approximately three-fold higher than that obtained when the LCQ instrument was used (130 peptide fragments). The individual components of the 26S proteasome were identified by 13.2 and 4.5 peptide fragments, 43.0% and 18.1% sequence coverage, and 821.9 and 292.4 protein scores on an average, using the LTQ and LCQ instruments, respectively. Twenty-nine proteins in 31 components were identified even by the LCQ instruments using our RP μLC system with high-resolution power, as shown in Fig. 1b. However, the peptide fragments of the remaining two components were not observed from the filtered database search results. A number

of peptide fragments belonging to the 26S proteasome (>700 peptides and >60% coverage on an average) were detected by an on-line 2-D SCX/RP μLC-MS/MS experiment with the same amount of digests using the LCQ instrument [10]. These results may simply indicate the difference in the scan speed for the limited, short analytical time between the LCQ and LTQ instruments and not a difference in the sensitivity. Accordingly, LC-MS/MS analysis using LTQ has a three-fold higher efficiency in identification in comparison with a conventional LCQ. This indicates that the LTQ has a superior protein identification capability. Thus, the introduction of LTQ into our μLC-MS/MS system resulted in a highly improved performance, in terms of both sensitivity and protein identification efficiency, for highly complex mixtures.

3.2 Human plasma proteome analysis

The usefulness and applicability of our automated protein profiling system coupled with LTQ for clinical proteomics have been examined by analyzing human plasma samples. In the course of clinical plasma proteome studies, we investigated two types of human plasma – one from healthy donors and the other from donors with lung adenocarcinoma. All plasma samples obtained from the three healthy (H-N, H-I, and H-S) and two adenocarcinoma (AC88 and AC94) cases were digested in a solution with trypsin after removing their HSA and IgG contents. The resultant peptide mixtures were diluted for the μLC-MS/MS analysis. Further, equivalent mixture samples from the three healthy donors (H-NIS; H-N:H-I:H-S, 1:1:1 in volume) and the two adenocarcinoma donors (AC8894; AC88:AC94, 1:1 in volume) were also prepared as average samples for each case. The individual and mixture samples (total seven samples) were analyzed by 1-D RP μLC-MS/MS analysis under analytical conditions for 90 min. The LTQ MS data was acquired by the double MS/MS method, in which the full MS acquisition is followed by two MS/MS scans of the two most intense precursor ions from the full MS scan. This is done to improve the protein identification results by the database search. Additionally, our established on-line 2-D SCX/RP μLC-MS/MS system using LTQ instead of the conventional LCQ instrument was also used to analyze the mixture samples (H-NIS and AC8894) along with an additional 1-D RP analysis. In the 2-D μLC-MS/MS analysis, six SCX separation runs were automatically carried out and analyzed by the RP μLC-MS/MS analysis under the same analytical conditions as described earlier. The total operation time for both 1-D and 2-D μLC-MS/MS analyses of a sample was within 11 h [10]. Figure 4 shows the base-peak chromatograms from 1-D (g) and 2-D

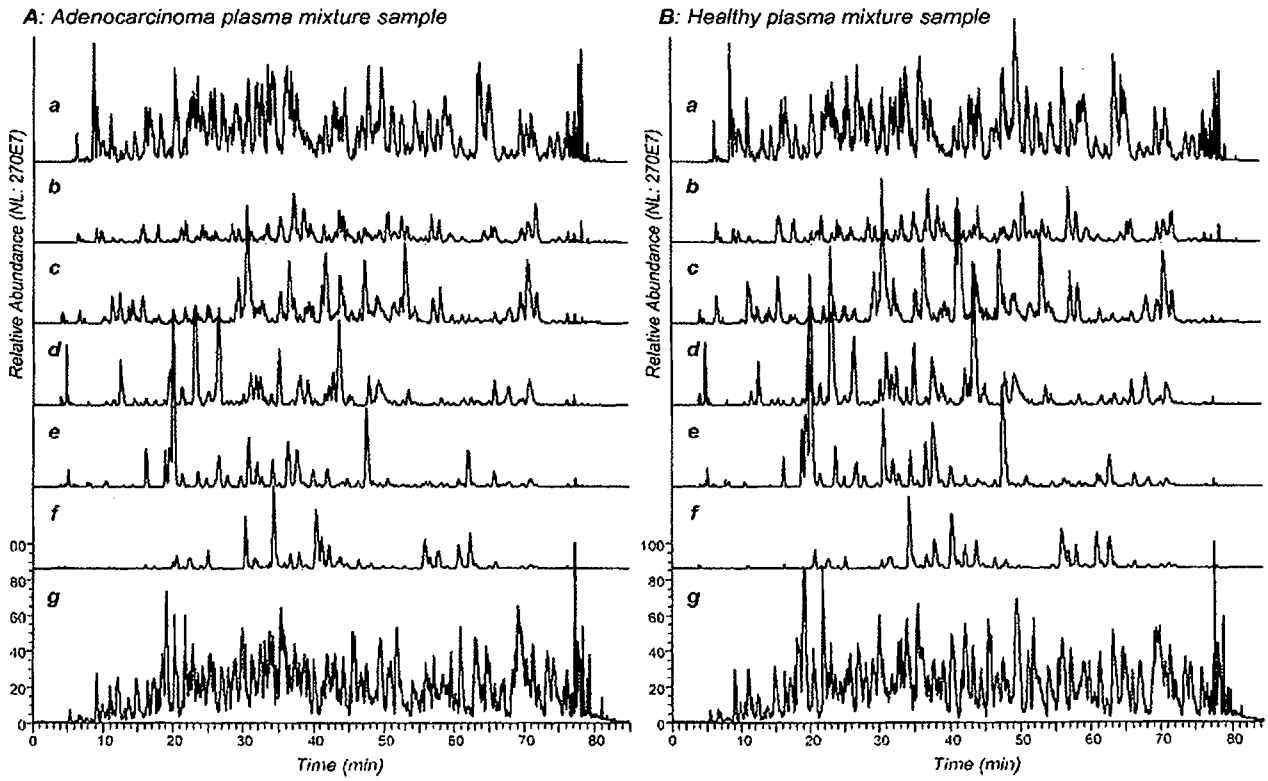

Figure 4. Base-peak chromatograms of the digested human plasma samples using 2-D SCX/RP (a–f) and 1-D RP μLC-MS/MS analyses (g). A, mixture of plasma digests of the healthy group; B, mixture of plasma digests of the adenocarcinoma group; a, 25 mM; b, 50 mM; c, 100 mM; d, 150 mM; e, 200 mM; f, 500 mM salt concentration SCX fractions for 2-D μLC-MS/MS analysis.

μLC-MS/MS analyses (a–f) of approximately 2 μg AID-HP tryptic digests corresponding to 0.4 μL original blood plasma sample. The 1-D RP μLC-MS/MS analysis provided approximately 10 000 MS/MS spectra for each sample, and the resultant data were evaluated using a Mascot database search against *H. sapien* subsets of the sequences in the Swiss-Prot database.

In order to achieve statistical confidence levels in identification of proteins from highly complex mixture samples, we investigated the thresholds as filters to extract data for the Mascot peptide score. We used the datasets, obtained from the 1-D μLC-MS/MS analysis with LTQ, of the digested 26S proteasome sample within our search tolerances. Since it is possible to identify several proteins from a single MS/MS spectrum based on the hit sequence varieties, the most significant peptide sequence that was ranked first (marked with bold red in the Mascot search results), which had the highest score among the hit varieties, was extracted from the entire datasets to prevent erroneous identifications and redundancy. The resulting peptide assignments were sorted according to their Mascot peptide score and intergraded into protein identification. Figure 5 (see also Supplementary Table A) shows the relationship between the peptide score ranges and either the number of peptide fragments assigned as 26S proteasome or other proteins identified erroneously. Although peptides with a score less than 50 were assigned to both the 26S proteasome and to the other proteins, all peptide fragments with a score of more than 50 were confidently identified as belonging to the 26S proteasome. For thresholds of peptide scores higher than 20, 30, and 40, the statistical identification confidence levels of the Mascot database search results were 70%, 90%, and 97%, respectively. Accordingly, we tentatively set peptide scores more than 30 and ranked first as the criterion for a broad protein identification index in order to integrate the datasets of the μLC-MS/MS analysis of plasma samples. Furthermore, to extract plasma proteins with a higher confidence level, we finally tried to apply the Swiss-Prot dataset (667 proteins) in a non-redundant list of 1175 distinct proteins that Anderson *et al.* have recently developed by combining four separate sources of human plasma proteome [3, 13–16].

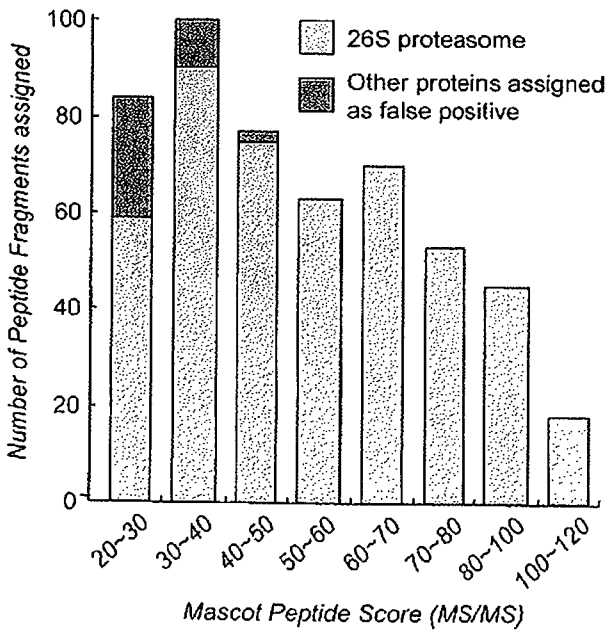

Figure 5. Mascot database search results of 1-D μLC-MS/MS analysis of the digested 26S proteasome. See also Supplementary Table A.

The Mascot database search results on plasma proteome analysis yielded data extracted under the thresholds of peptide scores higher than 30 and ranked first; an average of 108 proteins were detected from each 1-D μLC-MS/MS analysis. From the 2-D μLC-MS/MS analysis of the mixture sample of two groups, an average of 249 proteins was assigned as plasma proteome candidates. (Supplementary Table B). Additionally, entire datasets of these Mascot search results were integrated and processed with the data extraction. The results indicated that a total of 506 different proteins were listed, and 180 proteins were detected as common proteins from both groups (Fig. 6, in parentheses). Furthermore, plasma proteins with a high confidence level were extracted from these datasets using the Swiss-Prot dataset of the 667 plasma proteins reported. Figure 6 shows the diagrammatic representation of proteins found in the two groups comprising healthy individuals and adenocarcinoma patients, and the numbers are concordant with the proteins annotated as plasma proteins in 667 Swiss-Prot datasets. The results indicated that 84 and 85 proteins were extracted from the healthy and adenocarcinoma groups, respectively, and 69 proteins were common. In addition, 16 proteins were

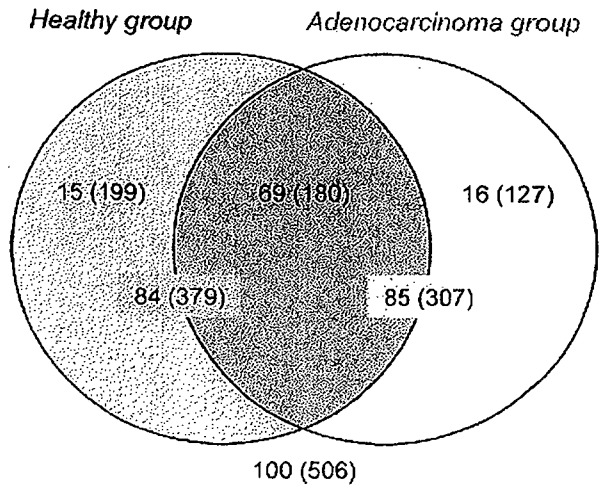

Figure 6. Diagrammatic representation of the proteins detected in the groups of healthy individuals and adenocarcinoma patients by 1-D and 2-D μLC-MS/MS analyses of their plasma samples. Numbers that belong to the peptide score thresholds higher than 30 served as the criteria for protein extraction (expressed in parentheses), and the numbers are concordant with the plasma proteins reported in the 667 Swiss-Prot datasets.

Table 2. Selected specific protein list of the healthy and lung adenocarcinoma groups in human blood plasma.

Accession No.	Protein name	Adenocarcinoma plasma samples					Healthy plasma samples					
		2-D analysis	1-D μLC-MS/MS analysis				2-D analysis	1-D μLC-MS/MS analysis				
		AC8894	AC88	AC94	AC8894-1	AC8894-2	H-NIS	H-N	H-I	H-S	H-NIS-1	H-NIS-2
P02649	Apolipoprotein E	A	C	A	A	A			C			
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	A	C	B	C		C		C	C	C	C
P04196	Histidine-rich glycoprotein	A										
P00748	Coagulation factor XII	B										
P00488	Coagulation factor XIII A chain	B					C					
P02570	Actin, cytoplasmic 1	C					A	A	C		C	
P27169	Serum paraoxonase/arylesterase 1						A					
P29312	14-3-3 protein zeta/delta								A			
P54108	Cysteine-rich secretory protein-3	C					B					
O14791	Apolipoprotein L1						B					
P02751	Fibronectin						B					
P06396	Gelsolin, plasma						B					

Two adenocarcinoma plasma samples (AC88 and AC94) and three healthy plasma samples (H-N, H-I, and H-S) were analyzed by 1-D μLC-MS/MS analysis. The mixture samples separated into both groups (AC8894 and H-NIS) were analyzed twice by 1-D and once by 2-D μLC-MS/MS analysis (2-D analysis). A–C indicated the presence of peptide fragment(s) assigned to the listed protein. A, mascot peptide score higher than 50; B, score 40 to 50; C, score 20 to 40. See also Supplementary Table B.

detected as specific significant proteins of the adenocarcinoma group, and 15 proteins were not detected in the adenocarcinoma group. Table 2 (see also Supplementary Table B) shows a list of the specific proteins assigned by peptide fragment(s) with scores higher than 40. These specific proteins detected from only one group might be candidate biomarkers of lung adenocarcinoma in human blood plasma. However, further statistical verification of our results through data accumulation of more disease plasma samples and the investigation concerning the reproducibility of protein identifications for each sample are necessary. Additionally, validation of these protein identifications by several biochemical approaches would be required. In the present study, we could indicate that several significant protein candidates in the plasma proteome are possibly associated with the pathological differences in lung adenocarcinoma. Functions of specific proteins and their correlations with adenocarcinoma along with the other proteins are not listed in this paper and will be reported elsewhere. These experimental achievements suggest that our automated 1-D and 2-D μLC-MS/MS protein profiling systems, in which the LTQ was incorporated, are powerful in identifying low-abundance proteins of great clinical importance, because these molecules directly report the occurrence and progress of various diseases.

4 Concluding remarks

In the course of the Human Plasma Organization Plasma Proteome Project, several research groups have prepared contrast reference specimens of human plasma using various technology platforms [17]. It is extremely important to catalog the plasma proteome as a protein database for clinical plasma proteomics. Applied technology platforms are very powerful, particularly for comprehensive broad protein identification of highly complex samples such as blood plasma. However, it seems difficult to stably and reproducibly apply these to routine clinical investigations that require a large number of proteome analysis runs for a large number of human samples. We have recently established a fully automated, high-throughput 2-D SCX/RP μLC-MS/MS protein profiling system, which can perform large-scale analysis in clinical proteomics [10]. In this study, the LTQ, which is superior to a conventional 3-D ITMS instrument in sensitivity and scan speed, was utilized in our high-throughput system, and it was evaluated by analyzing BSA and human 26S proteasome. Furthermore, the system was applied to plasma proteome analysis in a few cases of both healthy individuals and lung adenocarcinoma patients. The results confirmed that a 10-fold increase in terms of sensitivity was achieved in our system using the LTQ instrument for protein

identification. Further, in comparison with the conventional 3-D ITMS instrument, a three-fold higher number of peptide fragments was identified as belonging to the 26S proteasome, indicating significant improvement in resolution for the analytical time point. Additionally, approximately 250 and 100 different proteins were detected, based on the investigation criterion for a 90% confidence level of protein identification, from only 0.4 μL human plasma using 2-D and 1-D μLC-MS/MS analyses, respectively. The entire operation was automatically carried out within 11 h for both single 1-D and 2-D μLC-MS/MS analyses. From the protein identification datasets of both healthy and adenocarcinoma plasma samples, several disease-specific proteins were found in the human plasma based on the plasma proteome database reported earlier. Consequently, it was demonstrated that our μLC-MS/MS protein profiling system is feasible for large-scale analyses such as clinical plasma proteomics studies. Although plasma proteome analysis for clinical application still remains a great challenge due to the wide dynamic range of protein abundance, we shall continue further technological development of the large-scale proteome analysis based on the high-throughput μLC-MS/MS system reported in this paper. Such high-throughput and large-scale analysis of human plasma would lead to the discovery of new disease-associated protein markers with high sensitivity and high specificity in early disease detection and diagnosis, and this would revolutionize current therapeutics.

The authors gratefully acknowledge the technical assistance of Ms. Hisae Anyoji and Noriko Araki of Medical Proteoscope, Co. Inc., and medical doctors of the Department of Surgery, Tokyo Medical University, as well as the encouragement and support of AMR Inc., Tokyo, Japan. We are also deeply indebted to Drs. Hiroshi Matsumoto and Masayuki Kubota of Thermo Electron Co., Kanagawa, Japan for their assistance.

5 References

[1] Liotta, L. A., Ferrari, M., Petricoin, E., *Nature* 2003, *425*, 905.

[2] Anderson, N. L., Anderson, N. G., *Mol. Cell. Proteomics* 2002, *1*, 845–867.

[3] Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T. *et al.*, *Mol. Cell. Proteomics* 2004, *3*, 311–326.

[4] Washburn, M. P., Wolters, D., Yates, J. R. III, *Nat. Biotechnol.* 2001, *19*, 242–247.

[5] Wolters, D. A., Washburn, M. P., Yates, J. R., III, *Anal. Chem.* 2001, *73*, 5683–5690.

[6] Shen, Y., Jacobs, J. M., Camp, D. G., II, Fang, R. *et al.*, *Anal. Chem.* 2004, *76*, 1134–1144.

[7] Kawakami, T., Nagata, T., Muraguchi, A., Nishimura, T., *Electrophoresis* 2000, *21*, 1846–1852.

[8] Kawakami, T., Nagata, T., Muraguchi, A., Nishimura, T., *J. Chromatogr. B* 2003, *87*, 223–229.

[9] Fujii, K., Nakano, T., Kawamura, T., Usui, F. *et al.*, *J. Proteome Res.* 2004, *3*, 712–718.

[10] Fujii, K., Nakano, T., Hike, H., Usui, F. *et al.*, *J. Chromatogr. A* 2004, *1057*, 107–113.

[11] Tojo, H., *J. Chromatogr. A* 2004, *1056*, 223–228.

[12] http://www.matrixscience.com

[13] Pieper, R., Su, Q., Gatlin, C. L., Huang, S. T. *et al.*, *Proteomics* 2003, *3*, 422–432.

[14] Adkins, J. N., Varnum, S. M., Auberry, K. J., Moore, R. J. *et al.*, *Mol. Cell. Proteomics* 2002, *1*, 947–955.

[15] Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J. *et al.*, *Mol. Cell. Proteomics* 2003, *2*, 1096–1103.

[16] Pieper, R., Gatlin, C. L., Makusky, A. J., Russo, P. S. *et al.*, *Proteomics* 2003, *3*, 1345–1364.

[17] Omenn, G. S., *Proteomics* 2004, *4*, 1235–1240.

Development of a Novel Computer-Aided Diagnosis System for Automatic Discrimination of Malignant From Benign Solitary Pulmonary Nodules on Thin-Section Dynamic Computed Tomography

Kiyoshi Mori, MD, Noboru Niki, PhD,† Teturo Kondo, MD,* Yukari Kamiyama, MD,* Teturo Kodama, MD,* Yoshiki Kawada, PhD,† and Noriyuki Moriyama, MD‡*

Objectives: As an application of the computer-aided diagnosis of solitary pulmonary nodules (SPNs), 3-dimensional contrast-enhanced (CE) dynamic helical computed tomography (HCT) was performed to evaluate temporal changes in the internal structure of nodules to differentiate between benign nodules (BNs) and malignant nodules (MNs).

Methods: There were 62 SPNs (35 MNs and 27 BNs) included in this study. Scanning (2-mm collimation) was performed before and 2 and 4 minutes after CE dynamic HCT. The CT data were sent to a computer, and the pixels inside the nodule were characterized in terms of 3 parameters (attenuation, shape index, and curvedness value).

Results: Based on the CT data at 4 (MN: 1.81–27.1, BN: −42.8 to −3.29) minutes after CE–dynamic HCT, a score of 0 or higher can be assumed to indicate an MN.

Conclusions: Three-dimensional computer-aided diagnosis of the internal structure of SPNs using CE dynamic HCT was found to be effective for differentiating between BNs and MNs.

Key Words: coin lesion, pulmonary, computer-aided design, lung neoplasms, radiographic image enhancement, tomography, x-ray computed

(*J Comput Assist Tomogr* 2005;29:215–222)

The morphologic imaging diagnosis of solitary pulmonary nodules (SPNs) has been performed based on qualitative findings, mainly in computed tomography (CT) images, identified by diagnosticians when evaluating the characteristics of the nodule's margins, internal structure, and relations to surrounding structures.[1–3] The interpretation of these findings tends to differ, however, depending on the person performing the diagnosis, and diagnostic standards for differentiating between benign nodules (BNs) and malignant nodules (MNs) have yet to be established. The quantitative diagnosis of such lesions has been attempted based on the measurement of attenuation in the nodule. Attenuation has been used for the objective assessment of the internal structure of nodules and for the differential diagnosis of BNs and MNs.[4] There have also been reports on the use of contrast medium to evaluate changes in attenuation in nodules over time to differentiate between BNs and MNs.[5–8] In these studies, however, the attenuation in the nodules was measured in only a few slices. Moreover, because the region of interest (ROI) within the lesion was specified manually, the attenuation obtained showed a large degree of variation in different slices.

The use of helical scanning has facilitated the acquisition of volume data for the entire lesion, making it possible to analyze these image data using a computer.[9–11] In the present study, images of the entire lesion were obtained using contrast-enhanced (CE) dynamic helical computed tomography (HCT), and the changes in the density of the lesion over time were calculated with a computer and quantified in a 3-dimensional (3D) perspective for the differential diagnosis of BNs and MNs.

SUBJECTS AND METHODS

CT Imaging Conditions

Computed tomography images were obtained using an Xpress/SX system (Toshiba Corporation, Tokyo, Japan). The scanning parameters were a patient couch-top movement speed of 2 mm/s, a beam width of 2 mm, a tube voltage of 120 kV, a tube current of 200 mA, 1-second scanning, and an ROI of 200 mm. A total of 100 mL nonionic contrast medium (Iopamiron 300 Syringe; Nihon Schering, Tokyo, Japan) was injected at a rate of 2 mL/s using an autoinjector through a peripheral forearm vein. With the patient placed in the supine position and receiving supplemental oxygen via a nasal cannula (2 L/min), helical scanning covering the entire lesion (40–50 mm) was performed 3 times during breath-holding (before enhancement and 2 and 4 minutes after the start of

Received for publication August 10, 2004; accepted January 3, 2005.

From the *Department of Thoracic Diseases, Tochigi Cancer Center, Tochigi, Japan, †Department of Optical Science, University of Tokushima, Tokushima, Japan, and ‡Department of Radiology, National Cancer Center, Tokyo, Japan.

Supported in part by a grant-in aid for cancer research from the Ministry of Health and Welfare of Japan and the Second-Term Comprehensive 10-Year Strategy for Cancer Control.

Reprints: Kiyoshi Mori, Department of Thoracic Diseases, Tochigi Cancer Center, 4-9-13 Younan, Utsunomiya, Tochigi 320-0834, Japan (e-mail: kmori@tcc.pref.tochigi.jp).

contrast injection). Images were reconstructed at 1-mm intervals using a 180° algorithm.

Evaluation of CT Images

CT Image Processing

The tumor lesion was extracted from the thin-section CT images and then reconstructed to obtain a 3D CT image of the tumor.[9] The density was then calculated by characterizing the pixels inside the extracted nodule in terms of 3 parameters (attenuation, shape index, and curvedness value). Based on the calculated density values within each nodule before enhancement and 2 and 4 minutes after the start of injection, a linear discriminant function score was obtained for each time point.

Extraction of Lesions

After the reconstruction of 3D images from the thin-section CT images, ROIs containing the nodules were extracted from these images (Fig. 1). Using a deformable surface model, the nodules were then extracted from these ROIs.[9,12] In some cases, the nodules were located adjacent to the pleura. In such cases, preprocessing was required for extracting the region of the lung fields to separate the pleura from the nodules.[10] The segmentation of the nodule was based on a thresholding technique and selection of object connected components. In the present study, a method based on the deformable surface model proposed by Caselles et al[12] was used for extracting nodular regions with various density distributions.[10] This approach is based on deforming 3D surfaces, represented by level sets, toward the nodule boundary to be extracted in the 3D images. It automatically handles the changes in the surface topology during deformation. In this method, the nodular region is extracted by placing the initial curved surface within the nodule and then transforming the surface to conform to the margins of the nodule using a formula for the curved surface.[12] In this way, excessive overflow of blood vessels and bronchi relative to the curved surface can be prevented by adjusting the end points for curved surface transformation.[10]

Display and Assessment of Characteristic Values Within the Lesion

The pixels in the ROI, including the nodule, were expressed locally using a combination of 3 parameters: the attenuation, shape index, and curvedness value (Figs. 2, 3).[13,14]

The shape index and curvedness value are defined by the 3D curvature of the curved surface. The shape index ranges from 0 to 1. As the shape index approaches 0, the surface becomes increasingly convex (peak surface), and as the shape index approaches 1, the surface becomes more concave (pit surface). Thus, subtle curved surface structures can be expressed in numeric form. The curvedness value reflects the degree of curvature and ranges from 0 to 1. As the curvedness value approaches 0, the surface becomes flatter with less curvature. The concepts of the shape index and curvedness value can be easily understood when these 2 parameters are used for the curved surface of the tumor margins (boundary structures between the periphery of the nodule and the surrounding lung). These parameters, the shape index and curvedness value, obtained from the 3D curvature represent the concavoconvex structure of the curved surface and the degree of curvature of the curved surface, which are both determined from the relations between the target pixels and their adjacent pixels. These parameters can be regarded as indices of the uneven distribution of attenuation within the nodule. The histograms of the shape index, curvedness value, and attenuation within the nodule are obtained, and the scale of each histogram serves as a histogram characteristic value.[14] The Fisher linear discriminant classifier is commonly used in pattern classification and is an optimal classifier when the sample distributions are multivariate normal with equal covariance matrices.[15] The linear discriminant classifier was designed by using the histogram features. A leave-one-out procedure was performed to provide a less biased estimation of the performance of the linear discriminant classifier.[16] In this procedure, 1 nodule image is left out from the classifier design group and a linear discriminant function is formulated using the design group. The discriminant score is computed for the case that is left out by using the linear discriminant function obtained. This process cycles through the data set until every

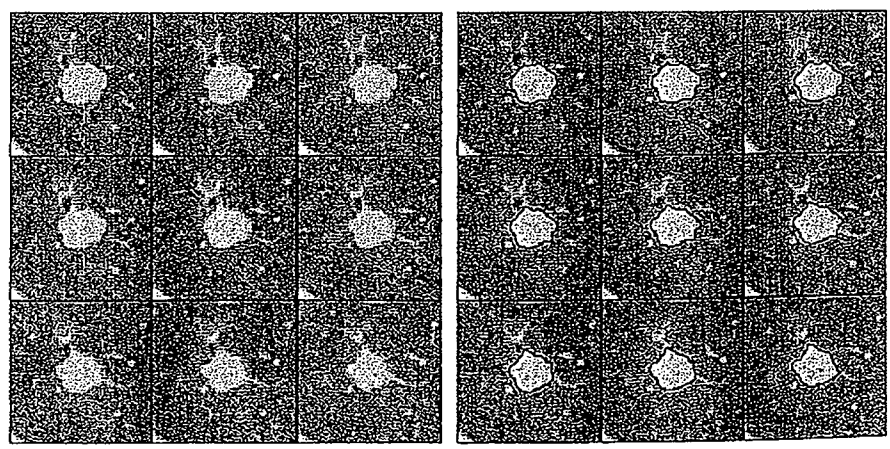

FIGURE 1. Extraction of a pulmonary nodule (case 8).

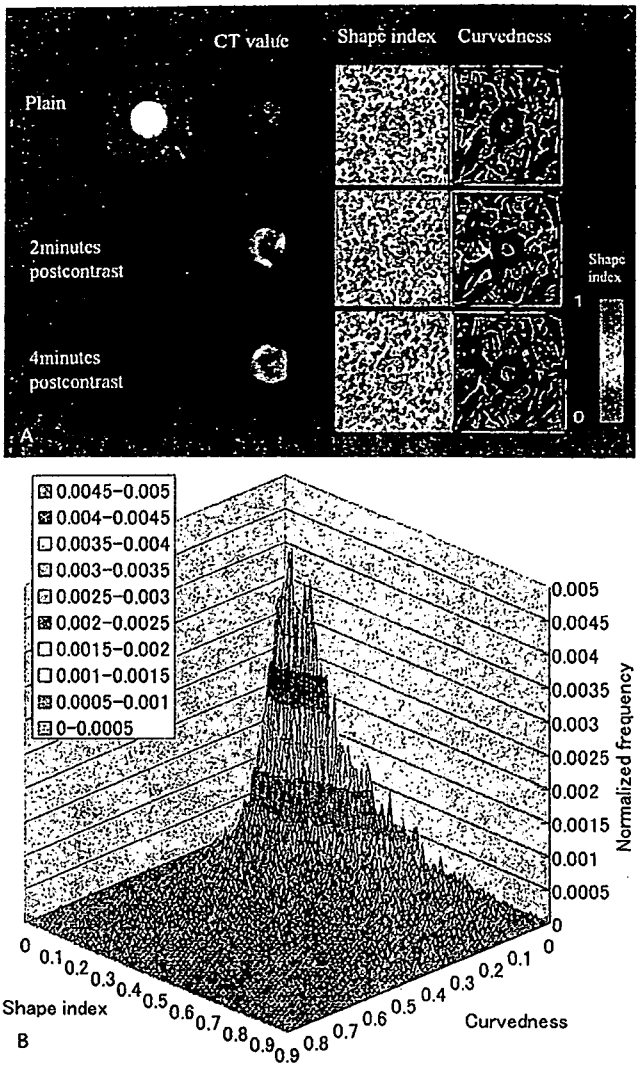

FIGURE 2. A, Characteristic values for a benign nodule (BN; case 37). B, Shape spectra showing a combination of the shape index and curvedness value 4 minutes after contrast enhancement. The *z* axis shows frequency. Most pixels inside a BN have a shape index close to 0 and a low curvedness value. This indicates that pixels inside a BN are mainly of the peak surface type with a smoothly curved surface.

nodule image is used. The overall evaluation time was approximately 4 minutes, including selection of the ROI from the CT images, extraction of the nodule, characterization of the pixels inside the nodule, and calculation of the linear discriminant function scores.

In the present study, each linear discriminant function score was computed from the shape index, curvedness value, and attenuation at each time point: before contrast enhancement and 2 and 4 minutes after the start of enhancement. The malignancy of SPNs was then retrospectively analyzed based on the scores obtained.

In addition, receiver operating characteristic curves were used to evaluate the effectiveness of the linear discriminant score in differentiating between BNs and MNs. Statistical significance was assessed using the unpaired Student *t* test.

Subjects

The subjects in this study were 72 consecutive patients who had undergone chest CT for the detailed examination of SPNs at our department from February 1998 to April 2000. They had only 1 target nodule by CT. Ten patients were not included in the assessment in this study, because CT images of the entire lesion could not be obtained over time (before contrast enhancement and 2 and 4 minutes after contrast enhancement) in these patients because of patient respiratory motion. The remaining 62 patients were evaluated. The mean

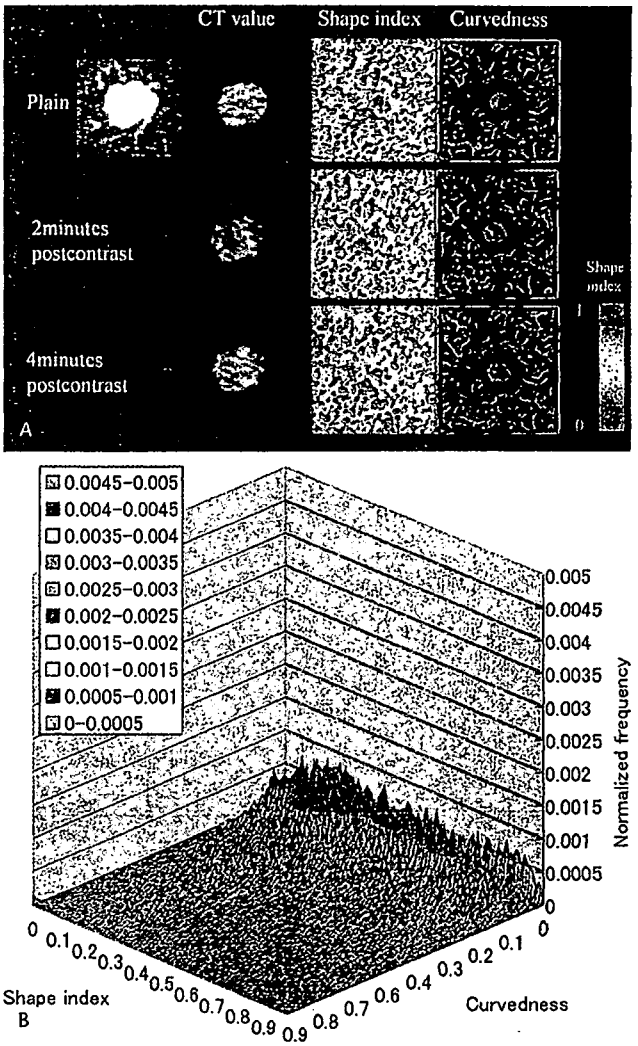

FIGURE 3. A, Characteristic values for a malignant nodule (MN; case 8). B, Shape spectra showing a combination of the shape index and curvedness value 4 minutes after contrast enhancement. Compared with a benign nodule (BN), pixels inside an MN show a wide distribution of shape index values, ranging from 0 to 1, and a high curvedness value. This indicates that pixels inside an MN tend to have pixels other than the peak surface type compared with a BN.