Table 1. Continued

Subunit	Phosphorylation sites	Amino acid sequence	*m/z*	Charge	Number of Supporting Information figures
Rpn13	Ser135	^{125}M$_{ox}$IGVLNNSSEpSDEEESNDEK^{144}	1161.4581	2	S3.50
Rpn13	Ser135,Ser140	^{125}MIGVLNNSSEpSDEEEpSNDEKQK^{146}	881.3414	3	S3.51
Rpn15	Ser12	Acetyl-^{2}STDVAAAQAQpSKIDLTK^{18}	934.9420	2	S1D

pS and pT represent phosphorylated Ser and Thr, respectively. pS* or pT* represent ambiguously identified sites. Mox and Cc represent oxidized Met and carbamidomethyl Cys, respectively. A MS spectrum of each peptide is shown in Supporting Information Figs. S1D and S3.

analyze the co- and post-translational modifications of the 26S proteasome purified from yeast.

4.1 *N*-Terminal modification

N-Terminal acetylation is the most common modification of proteins. In our previous study, we confirmed that the *N*-terminal Met is processed and the *N*-terminal amino group is modified on most 26S proteasome subunits according to the *N*-terminal processing and modification rule for eukaryotic proteins [33–35]. However, simultaneously, we found two exceptions to this rule in the yeast proteasome.

One exception is the free *N*-terminus of Rpt1, one of the six ATPase subunits of the 19S RP. In our previous study [9], the *N*-terminal modification of Rpt1 remained unclear. Edman degradation was performed to determine the *N*-terminal sequence of Rpt1, but it was not available from either the normal strain or *N*-acetyltransferase-deficient mutants. Therefore, at the beginning of this study, we predicted that the *N*-terminus of Rpt1 contained some modification group. However, no modification was found at the *N*-terminus of Rpt1 using MALDI-TOF/TOF-MS, and the *N*-terminal sequence was confirmed to be PPKEDWE. Usually, cleavage of the prolyl Pro bond is difficult in Edman degradation, resulting in carryover of a large amount of amino acids to the sequential degradation. This may be why the *N*-terminal Pro was not clearly detected in Edman degradation in our experiment. Moerschell *et al.* [34] suggested that Pro at the third position could inhibit the removal of Met from the penultimate residue, since Pro topologically distorts the *N*-terminus and interferes with the action of aminopeptidase. However, in the case of Rpt1, the initiator Met was removed, and the newly appeared *N*-terminus of Rpt1 was not modified, contrary to the rule.

Another exception is Rpt2, as described previously [9]. According to the *N*-terminal modification rule, it should be *N*-acetylated, but 50% of Rpt2 was found to be *N*-myristoylated, and the remaining 50% had a free *N*-terminus and no acetyl group. Recently, in mammals, *N*-myristoylation of Rpt2 was described for the human and murine cardiac proteasome [10, 18]. Similarly, in yeast, *N*-myristoylation is considered to be involved in protein–membrane interactions. Thus, in yeast, *N*-myristoylation of 50% of the Rpt2 molecules may be related to membrane binding of the proteasome.

In this study, we completed the identification of the *N*-terminal modifications of all 26S proteasome subunits, since we identified here the *N*-terminal modifications of Rpt1, Rpn13, and Rpn15.

4.2 Phosphorylation

Phosphorylation has been observed widely in proteasome complexes from many species and plays a role in regulating proteasome stability, assembly, and enzymatic activity [10–18] by altering the conformation of the complex [14]. In this study, we identified 33 Ser/Thr phosphorylation sites in 15 subunits in the yeast 26S proteasome.

To date, 63 Ser phosphorylation sites, 23 Thr phosphorylation sites, and 1 Tyr phosphorylation site have been identified (Table 2). Previously, we found that another Tyr residue was phosphorylated in the α7 subunit of the 20S proteasome [11], although its phosphorylation sites are unknown, revealing at least two Tyr phosphorylated subunits present in the 26S proteasome.

In *Caenorhabditis elegans*, interactions among proteasome subunits have been investigated using the yeast two-hybrid method [36]. Four potential interactions between two 19S ATPases and three 20S proteasome α-subunits (Rpt4/α2, α2/Rpt5, α4/Rpt4, and Rpt4/α7) have been predicted to be important in understanding the assembly of 19S RP and 20S proteasome, which is ATP-dependent and probably regulated by phosphorylation. In our studies, we found that highly phosphorylated subunits were arranged in the 19S RP base and the 20S proteasome α-subunits.

We previously reported that dephosphorylation of the 20S proteasome decreases protease/substrate affinity [11]. Several studies have shown that phosphorylation is related to enzymatic activities of the 20S proteasome and

association of 19S RP to the 20S proteasome through conformational changes [11, 14, 16, 17, 37]. However, the functional significance of the phosphorylation of 19S RP in yeast has not been determined completely. We considered the possibility of the involvement of phosphorylation in proteasome functions, such as substrate recognition, regulation of assembly and degradation, and ATPase activity. Here, we found that when 19S RP was dephosphorylated, the ATPase activity decreased (Fig. 1C), suggesting that the phosphorylation of 19S RP is involved in modulating its ATPase activity.

The ATPases of 19S RP are members of the AAA family, which possess weak ATPase activity [38]. The physiologically active form of these enzymes is often a homohexamer, and the ATP-binding site is positioned at the interface between subunits. Most AAA enzymes undergo conformational changes that are linked to ATP binding and hydrolysis, and these conformational changes are transmitted to substrate proteins [39]. We assume that a change in electric charge caused by the dephosphorylation of 19S RP topologically distorts the conformation, which interferes with the binding of ATP to the binding pockets of the ATPases, resulting in reduced ATPase activity.

4.3 Glycosylation

Several studies have reported on the glycosylation of the proteasome in some organisms [19–22], but not in yeast. First, we deduced that the 26S proteasome had no *N*-linked oligosaccharide chains because it consisted of nonsecreted proteins. In this study, this was confirmed by GelCode Glycoprotein Staining kit. Furthermore, we found that the 26S proteasome had no *O*-linked oligosaccharide chains by the same method.

We did, however, detect proteasome subunits that cross-reacted with the anti-*O*-GlcNAc antibody, although this cross-reaction was demonstrated to be nonspecific to these subunits, as described above. *O*-GlcNAcylation is found commonly in many multicellular eukaryotes. These organisms have *O*-GlcNAc transferase and *O*-GlcNAcase, which are involved in the regulation of *O*-GlcNAcylation. However, in yeast, no proteins with sequences homologous to these enzymes have been identified from the genome database [23, 40–43] and no *O*-GlcNAc transferase or *O*-GlcNAcase activity has been reported. From these facts, we concluded

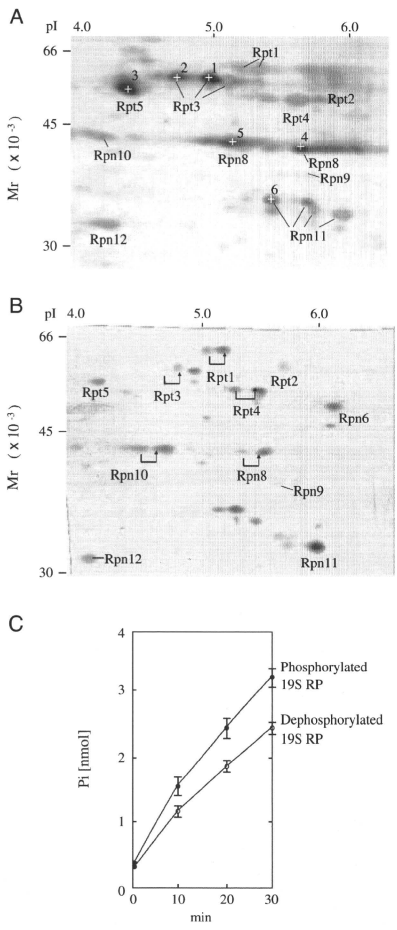

Figure 1. (A) 2-DE image of phosphorylated subunits of 19S RP stained with Pro-Q Diamond. The 19S RP (100 μg) was separated by 2-DE and stained with the Pro-Q Diamond phosphoprotein stain. After acquiring the fluorescence image, the gel was stained with CBB to visualize total proteins. In this figure, the fluorescence image (red) has been superimposed on the CBB image (blue). Spots (+) detected by the phosphostain were excised from the gel and identified by MS analysis after digestion with trypsin (Supporting Information Table S2). (B) Changes in electrophoretic mobility of the 19S RP subunits by 2-DE before and after dephosphorylation. Proteins separated by 2-DE were stained with CBB. The 2-DE pattern before dephosphorylation (red) has been superimposed on that after desphosphorylation (green). Spots indicated by arrows shifted on the 2-DE gels after dephosphorylation. Each spot was identified by MS/MS. (C) Changes in the amount of hydrolyzed phosphate released from the 19S RP before and after dephosphorylation.

Table 2. Overview of post-translational modifications in the yeast 26S proteasome

Subunit	Name	Accession	*N*-terminal mod.	Phosphorylation sites
α1	SCL1	PSA6_YEAST	*N*-acetyl Ser[a]	Ser237[b]
α2	PRE8	PSA2_YEAST	*N*-acetyl Thr[a]	Ser13[b], Ser53[b]
α3	PRE9	PSA4_YEAST	*N*-acetyl Gly[a]	
α4	PRE6	PSA7_YEAST	*N*-acetyl Ser[a]	**Thr60**[c], **Ser69**[b]
α5	PUP2	PSA5_YEAST	*N*-acetyl Met[a]	Ser16[d], Ser35[b], Thr55[e),f], **Ser56**[b),d),e),f], Ser87[b], **Ser251**[b),e]
α6	PRE5	PSA1_YEAST	*N*-acetyl Met[a]	Ser14[b),e], Thr16[b]
α7	PRE10	PSA3_YEAST	*N*-acetyl Thr[a]	Ser258[g], Ser263[g], Ser264[g], Thr278[b],Thr279[b]
β1	PRE3	PSB6_YEAST	Free[a]	Thr91[b]
β2	PUP1	PSB7_YEAST	Free[a]	
β3	PUP3	PSB3_YEAST	*N*-acetyl Ser[a]	Ser31[b], Ser33[f]
β4	PRE1	PSB2_YEAST	*N*-acetyl Met[a]	Ser76[b], Ser82[b]
β5	PRE2	PSB5_YEAST	Free[a]	Ser51[b]
β6	PRE7	PSB1_YEAST	Free[a]	Ser53[b], Ser209[c]
β7	PRE4	PSB4_YEAST	Free[a]	
Rpt1	RPT1	PRS7_YEAST	**Free**	**Ser164**[b),f], Ser231[b]
Rpt2	RPT2	PRS4_YEAST	*N*-myristoyl Gly[h]	Ser176[b),e], Thr178[b),e], Ser180[b),f], Tyr181[b),e], Ser182[b),e], Ser356[b], Thr357[b]
Rpt3	RPT3	PRS6B_YEAST	*N*-acetyl Met[h]	**Thr8**[b),d]
Rpt4	RPT4	PRS10_YEAST	*N*-acetyl Ser[h]	Ser292[b),f]
Rpt5	RPT5	PRS6A_YEAST	*N*-acetyl Ala[h]	**Thr11**, Thr27[b], Ser181[b),f], **Thr369**, **Thr370**
Rpt6	RPT6	PRS8_YEAST	*N*-acetyl Thr[h]	
Rpn1	RPN1	RPN1_YEAST	Free[a]	**Ser11**, Ser16[b], **Ser19**[b),f], **Thr24**[c),f], **Ser178**, **Ser179**, **Ser182**, **Ser187**[b], **Thr190**, **Thr640**, Ser695[b]
Rpn2	RPN2	RPN2_YEAST	*N*-acetyl Ser[a]	Ser56[b], **Thr801**[b], Ser883[b], Thr932[b]
Rpn3	RPN3	RPN3_YEAST	*N*-acetyl Ala[a]	Ser454[b]
Rpn5	RPN5	RPN5_YEAST	*N*-acetyl Ser[a]	**Ser26**
Rpn6	RPN6	RPN6_YEAST	*N*-acetyl Ser[a]	**Ser38**, Thr287[f]
Rpn7	RPN7	RPN7_YEAST	Free[a]	**Ser8**[b], Ser77[b]
Rpn8	RPN8	RPN8_YEAST	*N*-acetyl Ser[a]	**Ser44**[b], **Ser311**, **Ser314**[b),f),i], **Ser317**[b),e),f], **Ser319**[b),d),e),f),i], Ser321[b),f], **Thr327**[b]
Rpn9	RPN9	RPN9_YEAST	Free[a]	
Rpn10	RPN10	RPN10_YEAST	Free[a]	Ser259[f]
Rpn11	RPN11	RPN11_YEAST	*N*-acetyl Met[a]	**Ser243**, Thr262[b]
Rpn12	RPN12	RPN12_YEAST	Free[a]	
Rpn13	RPN13	RPN13_YEAST	***N*-acetyl Ser**	Ser132[b),e),f], Ser133[b),e], **Ser135**[b),e),f], **Ser140**[b),e]
Rpn15	SEM1	SEM1_YEAST	***N*-acetyl Ser**	**Ser12**[d]
BLM10	BLM10	BLM10_YEAST	–	Ser11[b], Ser29[b),c], Ser34[c], Ser35[c], **Ser56**[b),c),e),f], Ser62[b),c),f),i], Thr64[b),c),e),f),i], Thr66[b], Ser1041[b),c),d),e]

The modifications of Gothic amino acids were identified in this study. The modification of underlined amino acids was identified first in this study. References a) Kimura *et al.* [9]; b) Albuquerque *et al.* [50]; c) Chi *et al.* [48]; d) Gruhler *et al.* [46]; e) Li *et al.* [47]; f) Smolka *et al.* [49]; g) Iwafune *et al.* [12]; h) Kimura *et al.* [8]; i) Ficarro *et al.* [45].

that the 26S proteasome may contain no *O*-GlcNAcylated subunits.

4.4 Other post-translational modifications

So far, we have identified phosphorylation, N^{α}-acetylation, and N^{α}-myristoylation in the yeast 26S proteasome. Besides these post-translational modifications, oxidation of the proteasome has been reported. In the yeast 20S proteasome, Demasi *et al.* [44] reported that oxidation such as *S*-glutathionylation, affects chymotrypsin-like activity rather than trypsin-like activity, although the oxidation sites are unknown in the 20S proteasome.

Apart from yeast, Carrard *et al.* [3] reported that oxidation is related to the inhibition of protease activity of the mammal proteasome. Ishii *et al.* [6] showed that Rpt3 had extremely high carbonyl levels after oxidative stress induced by 15-deoxy-D-12,14-proteaglandin J_2 in human neuroblastoma cells, and that the oxidation of Rpt3 decreased the ATPase activity. Oxidation might play crucial roles in the regulation of the function of proteasome.

The authors thank Keiji Tanaka and Hikaru Ogura for invaluable suggestions and advice, and Hiroshi Kawasaki, Yuko Yamanaka, Yoko Ino and Ayuko Kimura for help in the experiments, and Daniel Finley for providing SDL66 cells. This

work was partially supported by Special Coordination Fund for Promoting Science and Technology "Creation and Innovation Centers for Advanced Interdisciplinary."

The authors have declared no conflict of interest.

5 References

[1] Coux, O., Tanaka, K., Goldberg, A. L., Structure and functions of the 20S and 26S proteasomes. *Annu. Rev. Biochem.* 1996, *65*, 801–847.

[2] Hershko, A., Ciechanover, A., The ubiquitin system. *Annu. Rev. Biochem.* 1998, *67*, 425–479.

[3] Carrard, G. C., Dieu, M., Raes, M., Toussaint, O. *et al.*, Impairment of proteasome structure and function in aging. *Int. J. Biochem. Cell. Biol.* 2002, *35*, 728–739.

[4] Chondrogianni, N., Gonos, E. S., Proteasome dysfunction in mammalian aging: steps and factors involved. *Exp. Gerontol.* 2005, *40*, 931–938.

[5] Ferrington, D. A., Kapphahn, R. J., Catalytic site-specific inhibition of the 20S proteasome by 4-hydroxynonenal. *FEBS Lett.* 2004, *578*, 217–223.

[6] Ishii, T., Sakurai, T., Usami, H., Uchida, K., Oxidative modification of proteasome: identification of an oxidation-sensitive subunit in 26S proteasome. *Biochemistry* 2005, *44*, 13893–13901.

[7] Gomes, A. V., Zong, C., Edmondson, R. D., Li, X. *et al.*, Mapping the murine cardiac 26S proteasome complexes. *Circ. Res.* 2006, *99*, 362–371.

[8] Kimura, Y., Takaoka, M., Tanaka, S., Sassa, H. *et al.*, N(alpha)-acetylation and proteolytic activity of the yeast 20S proteasome. *J. Biol. Chem.* 2000, *275*, 4635–4639.

[9] Kimura, Y., Saeki, Y., Yokosawa, H., Polevoda, B. *et al.*, N-Terminal modifications of the 19S regulatory particle subunits of the yeast proteasome. *Arch. Biochem. Biophys.* 2003, *409*, 341–348.

[10] Wang, X., Chen, C. F., Baker, P. R., Chen, P. L. *et al.*, Mass spectrometric characterization of the affinity-purified human 26S proteasome complex. *Biochemistry* 2007, *46*, 3553–3565.

[11] Iwafune, Y., Kawasaki, H., Hirano, H., Electrophoretic analysis of phosphorylation of the yeast 20S proteasome. *Electrophoresis* 2002, *23*, 329–338.

[12] Iwafune, Y., Kawasaki, H., Hirano, H., Identification of three phosphorylation sites in the alpha7 subunit of the yeast 20S proteasome *in vivo* using mass spectrometry. *Arch. Biochem. Biophys.* 2004, *431*, 9–15.

[13] Manson, G. G. F., Murray, R. Z., Pappin, D., Rivett, A. J., Phosphorylation of ATPase subunits of the 26S proteasome. *FEBS Lett.* 1998, *430*, 269–274.

[14] Manson, G. G. F., Hendil, K. B., Rivett, A. J., Phosphorylation of proteasomes in mammalian cells. Identification of two phosphorylated subunits and the effect of phosphorylation on activity. *Eur. J. Biochem.* 1996, *238*, 453–462.

[15] Murray, P. F., Pardo, P. S., Zelada, A. M., Passeron, S., *In vivo* and *in vitro* phosphorylation of *Candida albicans* 20S proteasome. *Arch. Biochem. Biophys.* 2002, *404*, 116–125.

[16] Satoh, K., Sasajima, H., Nyoumura, K., Yokosawa, H. *et al.*, Assembly of the 26S proteasome is regulated by phosphorylation of the p45/Rpt6 ATPase subunit. *Biochemistry* 2001, *40*, 314–319.

[17] Umeda, M., Manabe, Y., Uchimiya, H., Phosphorylation of the C2 subunit of the proteasome in rice (*Oryza sativa L.*). *FEBS Lett.* 1997, *403*, 313–317.

[18] Zong, C., Gomes, A. V., Drews, O., Li, X. *et al.*, Regulation of murine cardiac 20S proteasomes: role of associating partners. *Circ. Res.* 2006, *99*, 372–380.

[19] Schliephacke, M., Kremp, A., Schmid, H. P., Koehler, K. *et al.*, Prosomes (proteasomes) of higher plants. *Eur. J. Cell. Biol.* 1991, *55*, 114–121.

[20] Schmid, H. P., Vallon, R., Tomek, W., Kreutzer- Schmid, C. *et al.*, Glycosylation and deglycosylation of proteasomes (prosomes) from calf-liver cells: high abundance of neuraminic acid. *Biochimie* 1993, *75*, 905–910.

[21] Sümegi, M., Hunyadi-Gulyas, E., Medzihradszky, K. F., Udavardy, A., 26S proteasome subunits are *O*-linked *N*-acetylglucosamine-modified in *Drosophila melanogaster*. *Biochem. Biophys. Res. Commun.* 2003, *312*, 1284–1289.

[22] Zhang, F., Su, K., Yang, X., Bowe, D. B. *et al.*, *O*-GlcNAc modification is an endogenous inhibitor of the proteasome. *Cell* 2003, *115*, 715–725.

[23] Zachara, N. E., Hart, G. W., *O*-GlcNAc a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress. *Biochim. Biophys. Acta* 2004, *1673*, 13–28.

[24] Zachara, N. E., Hart, G. W., The emerging significance of *O*-GlcNAc in cellular regulation. *Chem. Rev.* 2002, *102*, 431–438.

[25] Zachara, N. E., O'Donnell, N., Cheung, W. D., Mercer, J. J. *et al.*, Dynamic *O*-GlcNAc modification of nucleocytoplasmic proteins in response to stress. A survival response of mammalian cells. *J. Biol. Chem.* 2004, *279*, 30133–30142.

[26] Saeki, Y., Toh-e, A., Yokosawa, H. Rapid isolation and characterization of the yeast proteasome regulatory complex. *Biochem. Biophys. Res. Commun.* 2000, *273*, 509–515.

[27] Leggett, D. S., Glickman, M. H., and Finley, D., Purification of proteasomes, proteasome subcomplexes, and proteasome-associated proteins from budding yeast. *Methods Mol. Biol.* 2005, *301*, 57–70.

[28] Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 1976, *7*, 248–254.

[29] Hirano, H., Varietal differences of leaf protein profiles in mulberry. *Phytochemistry* 1982, *21*, 1513–1518.

[30] Nabetani, T., Kim, Y. I., Watanabe, M., Ohashi, Y. *et al.*, Improved method of phosphopeptides enrichment using biphasic phosphate-binding tag/C18 tip for versatile analysis of phosphorylation dynamics. *Proteomics* 2009, *9*, 5525–5533.

[31] Ishihama, Y., Wei, F. Y., Aoshima, K., Sato, T. *et al.*, Enhancement of the efficiency of phosphoproteomic identification by removing phosphates after phosphopeptide enrichment. *J. Proteome Res.* 2007, *6*, 1139–1144.

[32] Hoffman, L., Rechsteiner, M., Nucleotidase activities of the 26S proteasome and its regulatory complex. *J. Biol. Chem.* 1996, *271*, 32538–32545.

[33] Polevoda, B., Sherman, F., *N*(alpha)-terminal acetylation of eukaryotic proteins. *J. Biol. Chem.* 2000, *275*, 36479–36482.

[34] Moerschell, R. P., Hosokawa, Y., Tsunasawa, S., Scherman, F., The specificities of yeast methionine aminopeptidase and acetylation of amino-terminal methionine *in vivo*. Processing of altered iso-1-cytochromes c created by oligonucleotide transformation. *J. Biol. Chem.* 1990, *265*, 19638–19643.

[35] Polevoda, B., Sherman, F., *N*-terminal acetyltransferases and sequence requirements for *N*-terminal acetylation of eukaryotic proteins. *J. Mol. Biol.* 2003, *325*, 595–622.

[36] Davy, A., Bello, P., Thierry-Mieg, N., Vaglio, P. *et al.*, Protein–protein interaction map of the *Caenorhabditis elegans* 26S proteasome. *EMBO Rep.* 2001, *2*, 821–828.

[37] Rivett, A., Bose, S., Brooks, P., Broadfoot, K., Regulation of proteasome complexes by gamma-interferon and phosphorylation. *Biochimie* 2001, *83*, 363–366.

[38] Song, C., Wang, Q., Li, C. C. H., ATPase activity of p97-valosin-containing protein (VCP). D2 mediates the major enzyme activity, and D1 contributes to the heat-induced activity. *J. Biol. Chem.* 2003, *278*, 3648–3655.

[39] Hanson, P. I., Whiteheart, S. W., AAA+proteins: have engine, will work. *Nat. Rev. Mol. Cell. Biol.* 2005, *6*, 519–529.

[40] Comer, F. I., Hart, G. W., *O*-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between *O*-GlcNAc and *O*-phosphate. *J. Biol. Chem.* 2000, *275*, 29179–29182.

[41] Hanover, J. A., Glycan-dependent signaling: *O*-linked *N*-acetylglucosamine. *FASEB J.* 2001, *11*, 1865–1876.

[42] Kreppel, L. K., Blomberg, M. A., Hart, G. W., Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *J. Biol. Chem.* 1997, *272*, 9308–9315.

[43] Schirm, M., Kalmokoff, M., Aubry, A., Thibault, P. *et al.*, Flagellin from *Listeria monocytogenes* is glycosylated with beta-*O*-linked *N*-acetylglucosamine. *J. Bacteriol.* 2004, *186*, 6721–6727.

[44] Demasi, M., Silva, G. M., Netto, L. E. S., 20S proteasome from *Saccharomyces cerevisiae* is responsive to redox modifications and is *S*-glutathionylated. *J. Biol. Chem.* 2003, *278*, 679–685.

[45] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J. *et al.*, Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 2002, *20*, 301–305.

[46] Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P. *et al.*, Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* 2005, *4*, 310–327.

[47] Li, X., Gerber, S. A., Rudner, A. D., Beausoleil, S. A. *et al.*, Large-scale phosphorylation analysis of alpha-factor-arrested *Saccharomyces cerevisiae*. *J. Proteome Res.* 2007, *6*, 1190–1197.

[48] Chi, A., Huttenhower, C., Geer, L. Y., Coon, J. J. *et al.* Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 2193–2198.

[49] Smolka, M. B., Albuquerque, C. P., Chen, S.-H., Zhou, H., Proteome-wide identification of *in vivo* targets of DNA damage checkpoint kinases. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 10364–10369.

[50] Albuquerque, C. P., Smolka, M. B., Payne, S. H., Bafna, V. *et al.*, A multidimensional chromatography technology for in-depth phosphoproteome analysis. *Mol. Cell. Proteomics* 2008, *7*, 1389–1396.

DOI 10.1002/pmic.201000349 *Proteomics* 2010, *10*, 3884–3895

RESEARCH ARTICLE

Characterization of multiple alternative forms of heterogeneous nuclear ribonucleoprotein K by phosphate-affinity electrophoresis

Yayoi Kimura[1,2], Kayoko Nagata[1], Nobutake Suzuki[1], Ryo Yokoyama[1], Yuko Yamanaka[2], Hiroshi Kitamura[1,3], Hisashi Hirano[2] and Osamu Ohara[1,4]

[1] Laboratory for Immunogenomics, RIKEN Research Center for Allergy and Immunology, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, Japan
[2] Graduate School of Nanobioscience, Yokohama City University, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, Japan
[3] Department of Comparative and Experimental Medicine, Graduate School of Medical Sciences, Nagoya City University, Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya, Japan
[4] Department of Human Genome Research, Kazusa DNA Research Institute, Kazusa-kamatari, Kisarazu, Chiba, Japan

Received: June 2, 2010
Revised: August 13, 2010
Accepted: August 22, 2010

The phosphorylation of heterogeneous nuclear ribonucleoprotein K (hnRNP K) is thought to play an important role in cell regulation and signal transduction. However, the relationship between hnRNP K phosphorylation and cellular events has only been indirectly examined, and the phosphorylated forms of endogenous hnRNP K have not been biochemically characterized in detail. In this study, we extensively examined the phosphorylated forms of endogenous hnRNP K by direct protein–chemical characterization using phosphate-affinity electrophoresis followed by immunoblotting and MS. Phosphate-affinity electrophoresis enabled us to sensitively detect and separate the phosphorylated forms of hnRNP K. When we used 2-DE with phosphate-affinity SDS-PAGE in the second dimension, the nuclear fraction contained more than 20 spots of endogenous hnRNP K on the 2-D map. We determined that the multiple forms of hnRNP K were produced mainly by alternative splicing of the single hnRNP K gene and phosphorylation of Ser116 and/or Ser284. Furthermore, the subcellular localization of these proteins revealed by the 2-D gel correlated with their phosphorylation states and alternative splicing patterns. The results also indicated that the multiple forms of hnRNP K were differentially modulated in response to external stimulation with bacterial lipopolysaccharide or serum.

Keywords:
2-D phosphate-affinity gel electrophoresis / Alternative splicing / Animal Proteomics / Heterogeneous nuclear ribonucleoprotein K / Mass spectrometry / Phosphorylation

1 Introduction

Post-translational modifications (PTMs) frequently play a key role in regulating protein function. Thus, quantitative and qualitative information about PTMs is crucial to understand the biological functions, subcellular localization and intracellular molecular interactions of many proteins [1, 2]. Although the current MS techniques have provided us with powerful approaches to analyze PTMs [3], it is still challenging to quantitatively characterize all the PTMs that occur in a particular protein. Furthermore, although MS techniques enable us to quantitatively analyze PTMs associated with particular amino acid residue(s), these techniques alone cannot determine the combination of PTMs that are associated with an entire protein. In this regard,

Correspondence: Dr. Osamu Ohara, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
E-mail: oosamu@rcai.riken.jp
Fax: +81-45-503-9694

Abbreviations: LPS, lipopolysaccharide; **PTM,** Post-translational modification

PTM-dependent protein separation is expected to play a pivotal role in elucidating the relationship between PTMs and protein function.

Protein phosphorylation is the most common PTM and therefore one of the most important PTMs. In many cases, a protein can be reversibly phosphorylated at multiple sites, and these modifications can either act independently or synergistically when they occur simultaneously [4]. Therefore, it is important to investigate both the phosphorylation states of individual sites in a given protein as well as the overall phosphorylation state of the protein to understand protein function. 2-DE is a powerful analytical method that uses orthogonal separation by p*I* and *M*r, and this technique can be used to simultaneously separate and isolate similar but distinct protein variants produced by PTMs, such as phosphorylation, glycosylation and sulfation [5, 6]. Recently, a new technique was developed to analyze phosphoproteins using the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand as a copolymer in separating gel matrix in SDS-PAGE or 2-DE [7–9]. This phosphate-affinity electrophoresis enables phosphoproteins to be separated from their non-phosphoprotein counterparts based on their phosphorylation states.

Heterogeneous nuclear ribonucleoprotein (hnRNP) K is an abundant RNA-binding protein that is found in multiple subcellular compartments, including the nucleus, cytoplasm, and mitochondria [10, 11]. The hnRNP K protein has an *N*-terminal bipartite nuclear localization signal and an hnRNP K nuclear shuttling domain located immediately before the C-terminal KH3 domain [10]. The hnRNP K protein is involved in various cellular processes, such as chromatin remodeling, transcription, mRNA processing, and translation [12, 13] and interacts with a wide range of binding partners through multiple domains. Many of these interactions have been shown to be regulated by phosphorylation cascades that are induced by extracellular signals [13, 14].

Previous reports have shown that hnRNP K undergoes several PTMs, including phosphorylation, methylation, and sumoylation [15–21]. For example, c-Src-mediated tyrosine phosphorylation of hnRNP K results in the inhibition of mRNA translation [19]. Extracellular-signal-regulated kinase (ERK)-dependent phosphorylation of hnRNP K on Ser284 and Ser353 has been shown to result in the cytoplasmic accumulation of hnRNP K and enhanced translational repression [17]. HnRNP K is also phosphorylated by C-Jun N-terminal kinase (JNK) Ser216 and Ser353, which contributes to hnRNP K-mediated transcriptional activities [18]. Interleukin-1-responsive kinase and protein kinase Cδ have been shown to phosphorylate hnRNP K *in vitro* [15, 16]. Overall, these reports clearly demonstrate that phosphorylation critically governs the cellular functions of hnRNP K. However, the multiple phosphorylated forms of hnRNP K have not been extensively examined using protein–chemical characterization.

Therefore, in this study, we examined the multiple forms of endogenous hnRNP K using phosphate-affinity electrophoresis coupled with MS analysis. Consequently, we showed that unbiased phosphorylation of various Ser residues and alternative splicing generate multiple forms of endogenous hnRNP K. In addition, we determined that these modifications depend on external stimulation with bacterial lipopolysaccharide (LPS) or serum, which modulated both the function and subcellular localization of these modified proteins.

2 Materials and methods

2.1 Cells and treatment

Mouse macrophage-like J774.1 cells (RIKEN Bioresource Center, Tsukuba, Japan) were maintained in RPMI 1640 supplemented with 10% FBS. J774.1 cells were seeded at a density of 2.0×10^5 cells/mL two days prior to harvest, and then stimulated with LPS (*Escherichia coli* 055:B5; Sigma-Aldrich, St. Louis, MO, USA) at a final concentration of 100 ng/mL for 2 or 4 h. To examine serum-induced phosphorylation of hnRNP K, HeLa cells were first cultured in DMEM supplemented with 1% serum for 20 h, and then treated with DMEM containing 10% serum for 30 min at 80% confluence.

2.2 Subcellular fractionation and protein preparation

Harvested cells were resuspended in an equal volume of ice-cold cytoplasmic extraction buffer [10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 10 mM $MgCl_2$, 0.34 M sucrose, Complete protease inhibitor EDTA free (Roche Diagnostics, Basel, Switzerland), 1% v/v Phosphatase inhibitor cocktails 1 and 2 (P2850 and P5726, respectively; Sigma-Aldrich), and 0.1% v/v Carbobenzoxy-Leu-Leu-Leu-H (aldehyde) (MG132; Peptide Institute, Osaka, Japan)] containing 0.5% v/v NP-40. After incubating on ice for 2.5 min, the cytoplasmic fraction was recovered by centrifuging at $4000 \times g$ for 2 min at 4°C. After washing twice with 1 mL ice-cold cytoplasmic extraction buffer containing 0.25% v/v NP-40, the nuclear pellet was resuspended in an equal volume of nuclear extraction buffer [10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 10% v/v glycerol, Complete protease inhibitor EDTA free, 1% v/v Phosphatase inhibitor cocktails 1 and 2, and 0.1% v/v MG132], and then diluted twofold with nuclear extraction buffer containing 500 mM NaCl. After mixing, the nuclear fraction was clarified by centrifuging at $17\,000 \times g$ for 10 min. To prepare whole cell extracts, the cells were lysed with whole cell lysis buffer [7 M urea, 2 M thiourea, 4% w/v CHAPS, and 2 mM Tris(2-carboxyethyl)phosphine] or SDS-PAGE sample buffer. Then, the whole cell extracts were clarified by centrifuging at $17\,000 \times g$ for 10 min. Each isolated fraction was stored as −80°C until further use. The protein concentration of each

sample was determined with a Bio-Rad Protein Assay kit (Bio-Rad Laboratories, Hercules, CA, USA) using bovine immunoglobulin (Bio-Rad Laboratories) as a protein standard [22]. For immunoprecipitation, the nuclear or cytoplasmic fractions were diluted with a two or tenfold volume of PBS, respectively, and then incubated with an anti-hnRNP K/J antibody (Santa Cruz Biotechnology, CA, USA) bound to protein A–Sepharose (GE Healthcare, Little Calfont, UK) for 3 h at 4°C. After washing the protein A–Sepharose–antibody complexes, the immunoprecipitated hnRNP K proteins were eluted with 2-DE rehydration buffer [7 M urea, 2 M thiourea, 4% w/v CHAPS, 20 mM DTT, and 0.5% v/v IEF gel buffer, pH 4–7 (GE Healthcare)].

2.3 Protein dephosphorylation *in vitro*

For lambda protein phosphatase treatment (λ-PPase; New England Biolabs, MA, USA), phosphatase inhibitor cocktails were omitted from the cytoplasmic and nuclear extraction buffer. A protein extract (2 μg/μL) was incubated at 30°C for 5 h in a reaction buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT, 0.1 mM EGTA, 0.01% Brij35, and Complete protease inhibitor EDTA free) containing 16 U λ-PPase *per* 10 μg protein. The reaction was stopped by adding either SDS-PAGE sample buffer or 2-DE rehydration buffer.

2.4 Protein electrophoresis

Conventional SDS-PAGE was performed using polyacrylamide slab gels purchased from BIOCRAFT (Tokyo, Japan). Phosphate-affinity SDS-PAGE with 25 μM polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand (Phos-tag AAL-107; Phos-tag Consortium, Hiroshima, Japan) was performed according to the manufacturer's instructions. Large 2-DE (IEF/SDS-PAGE; 170 × 180 × 1 mm) was carried out according to the method previously described [23]. When a mini 2-D gel (60 × 80 × 1 mm) was used, 2-DE was performed as described below. Protein (10 μg) was diluted in 2-DE rehydration buffer (final volume of 155 μL) and then applied to an IPG dry strip gel (7 cm, pH 4.7–5.9; Bio-Rad Laboratories) by in-gel rehydration for 16 h at room temperature. IEF in the first dimension was performed using the Zoom IPG Runner system (Invitrogen, San Diego, CA, USA) according to the manufacturer's instructions. For the second dimension, a home-made 7.5% w/v polyacrylamide gel with or without 25 μM polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand was used. After electrophoresis, the separated proteins were stained with SYPRO Ruby protein gel stain and/or ProQ Diamond phosphoprotein gel stain (Molecular Probes, Eugene, OR, USA) or were subjected to immunoblotting analysis on polyvinylidene difluoride membranes.

2.5 Immunoblotting analysis

The following antibodies were used for immunoblotting: anti-GAPDH, anti-cytochrome *c*, anti-NFκB p65, anti-hnRNP K/J (Santa Cruz Biotechnology), anti-HDAC1, anti-p44/42 MAPK (Erk1/2), anti-phospho-p44/42 MAPK (Thr202/Tyr204) (p-Erk1/2) (Cell Signaling Technology, MA, USA), and anti-phosphotyrosine (clone PY99, Santa Cruz Biotechnology; clone 4G10, Cell Signaling Technology). These primary antibodies were visualized with HRP-conjugated secondary antibodies using an ECL plus Western Blotting Detection System (GE Healthcare) on a luminescence image analyzer (LAS-1000 mini; Fuji Film, Tokyo, Japan).

2.6 In-gel digestion and MS analysis

The protein spots were excised from SYPRO Ruby-stained 2-DE gels, reduced with 25 mM NH_4HCO_3 containing 10 mM DTT, and then alkylated with 25 mM NH_4HCO_3 containing 55 mM iodoacetamide in the gel. Next, the gel pieces were completely dried and reconstituted with 25 mM NH_4HCO_3 containing 0.1% v/v acid-labile surfactant (RapiGest; Waters, MA, USA) for 30 min at room temperature. After the gel pieces were completely dried, in-gel digestion was performed using the following conditions: 60 ng trypsin (Promega, Madison, WI, USA) in 25 mM NH_4HCO_3 at 30°C overnight, 60 ng endoproteinase LysC (Wako Pure Chemical, Osaka, Japan) in 25 mM NH_4HCO_3 at 37°C overnight, 60 ng endoproteinase LysC in 25 mM NH_4HCO_3 at 37°C overnight followed by 60 ng trypsin at 37°C overnight, 60 ng endoproteinase GluC (Sigma) in 25 mM NH_4HCO_3 at 30°C overnight followed by 60 ng endoproteinase LysC at 30°C overnight, or 60 ng endoproteinase AspN (Roche Diagnostics) in 25 mM NH_4HCO_3 at 37°C overnight. To identify C-terminal peptides and other peptides, the resulting peptides were desalted using ZipTip-C18μ (Millipore, Bedford, MA, USA) and then eluted directly with matrix solution containing 4 mg/mL CHCA (Sigma-Aldrich) on a target plate. For phosphopeptide analysis, the resulting peptides were extracted from the gel piece with 0.1% v/v TFA and 50% v/v ACN followed by 0.067% v/v TFA and 67% v/v ACN. The phosphopeptides were enriched with the Phos-trap phosphopeptide enrichment kit (PerkinElmer, Boston, MA, USA) following the manufacturer's instructions with minor modifications. Briefly, peptides bound to the TiO_2 beads were eluted twice with 4 μL of the elution buffer provided in the kit and then the eluates were combined and concentrated. The concentrated peptides and matrix solution (0.6 μL of each) were mixed on a target plate. These peptide samples were analyzed on an ABI4800 proteomics analyzer (Applied Biosystems, Foster City, CA, USA) using the 4000 Series Explorer Software version 3.5.1. The obtained MS and MS/MS data were searched against the IPI database (version

3.57; 56 091 mouse protein sequences) using the MASCOT program, version 2.2.01 (Matrix Science, London, UK) to identify proteins or PTMs. The search parameters were as follows: each proteinase digestion with two or three missed cleavage permitted, variable modifications (*N*-terminal acetylation, oxidation of methionine, carbamidomethylation of cysteine, dimethylation of arginine, and phosphorylation of serine or threonine), peptide mass tolerance for MS data ±100 ppm, and fragment mass tolerance ±0.3 Da. Search results that yielded a peptide match score that was greater than or equal to the individual peptide score identity threshold ($p<0.05$) were considered positive identifications.

2.7 Plasmids and cell-free protein synthesis

The protein-coding region of mouse hnRNP K cDNA was PCR amplified and cloned into a pGEM®-T easy vector system (Promega). The clones were confirmed to have a nucleotide sequence that was identical to that of hnRNP K in the public databases (AK011428 and AK150848 of isoforms 1 and 2, respectively). For cell-free protein synthesis, hnRNP K cDNA was subcloned into pF3AHT or pF1K T7 (Promega). Mutations were introduced into the serine phosphorylation sites (S116A, S284A, and S353A) of hnRNP K using a Quick change site-directed mutagenesis kit (Stratagene, La Jolla, CA, USA) and confirmed by DNA sequencing. Cell-free transcription and translation were performed using a wheat germ extract (TNT SP6 High-Yield Protein Expression System; Promega). *In vitro* translation was conducted at 25°C for 2 h according to the manufacturer's instructions. The translation reactions were stopped by adding 2-DE rehydration buffer.

3 Results

3.1 Detection of phosphorylated hnRNP K forms that are modulated by LPS stimulation in J774.1 cells

In order to identify proteins that are modulated upon LPS stimulation, we previously compared the 2-DE protein profiles of J774.1 macrophage-like cells that were incubated in the presence and absence of LPS (100 ng/mL) for 2 h [23]. A comparison of SYPRO Ruby-stained gel images revealed that the intensities of four spots (a, c, d, and f) among six spots (a–f) of hnRNP K were reproducibly modulated by LPS stimulation (Fig. 1A). On the other hand, the gel images obtained by Pro-Q Diamond phosphoprotein fluorescence dye staining suggested that four spots (a, b, d, and e) and two major spots (c and f) represented the phosphorylated and non-phosphorylated forms of hnRNP K, respectively. These results suggested that the phosphorylation states of hnRNP K in J774.1 cells were considerably affected by LPS stimulation (Fig. 1A). To further characterize the LPS-dependent phosphorylation of endogenous hnRNP K in J774.1 cells, we employed phosphate-affinity SDS-PAGE using the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand followed by immunoblotting with an anti-hnRNP K/J antibody. The Phos-tag™ ligand is a phosphate-binding compound that can be incorporated into polyacrylamide gels and results in a mobility shift in phosphorylated forms that correlates with the degree of phosphorylation [7, 8]. Therefore, phosphate-affinity SDS-PAGE followed by immunoblotting enabled us to readily compare the phosphorylation states of endogenous hnRNP K. In a practical experimental setting, we separately analyzed whole cell lysate and cytoplasmic/nuclear fractions using phosphate-affinity SDS-PAGE (Fig. 1B and C). Consequently, we found that phosphate-affinity SDS-PAGE resolved hnRNP K in the cytoplasmic and whole cell/nuclear fractions into two and five distinct bands, respectively. However, conventional SDS-PAGE analysis of hnRNP K in the cytoplasmic and whole cell/nuclear fractions resulted in a single band of 66 kDa and double bands of 66 and 64 kDa, respectively (Fig. 1C). Because some bands on the phosphate-affinity SDS-PAGE analysis disappeared after the lysates were treated with λ-PPase, the observed differences in electrophoretic mobility were attributed to phosphorylation (Fig. 1D). Thus, these results indicated that several distinct phosphorylated forms of hnRNP K were modulated in response to LPS stimulation and differentially localized in J774.1 cells without any significant change in total expression level (Fig. 1C).

To more precisely characterize the alternative forms of endogenous hnRNP K in J774.1 cells, we used phosphate-affinity 2-DE, where a narrow-range IPG gel and phosphate-affinity SDS-PAGE were applied in the first and second dimensions, respectively. After the phosphate-affinity 2-DE, hnRNP K was specifically detected by immunoblotting with an anti-hnRNP K/J antibody. We expected that the inclusion of IPG would enable us to detect subtle changes in the electrophoretic characteristics of hnRNP K caused by PTMs [5]. In fact, phosphate-affinity 2-DE clearly yielded a higher resolution of multiple hnRNP K forms and revealed that endogenous hnRNP K in the cytoplasmic and nuclear fractions was composed of at least 4 and 16 multiple forms with distinct p*I* and phosphorylation states, respectively (Fig. 2A). Moreover, the results indicated that the cytoplasmic and the nuclear fractions were distinct in terms of the contents of the multiple forms and that their contents changed characteristically in response to LPS stimulation (Fig. 2A).

3.2 Structural characterization and subcellular localization of multiple forms of endogenous hnRNP K

Next, we structurally characterized the multiple forms of hnRNP K that were detected by phosphate-affinity 2-DE in J774.1 cells. The multiple forms of hnRNP K thus identified (20 spots, Fig. 3A) were structurally characterized as described below.

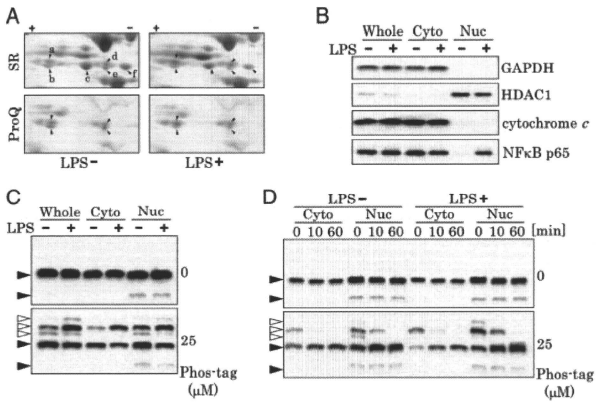

Figure 1. HnRNP K forms in whole-cell, cytoplasmic and nuclear fractions of LPS-stimulated J774.1 cells. (A) An expanded 2-DE gel image of the endogenous hnRNP K detection region is shown. Whole cell extracts [whole cell lysis buffer] of J774.1 cells were separated with large conventional 2-DE (IEF/SDS-PAGE) followed by SYPRO Ruby staining (SR) or Pro-Q Diamond staining (ProQ). Six spots (a–f) indicated multiple spots of hnRNP K. (B) To confirm the fractionation and LPS stimulation, conventional 10% SDS-PAGE followed by immunoblotting was performed using an anti-GAPDH antibody as a cytoplasmic marker, an anti-HDAC1 antibody as a nuclear marker, an anti-cytochrome *c* antibody as a mitochondrial marker, and an anti-NFκB p65 antibody as a marker of LPS stimulation. (C and D) To analyze the phosphorylation states of hnRNP K, 7.5% w/v SDS-PAGE in the absence (upper panel) and presence (bottom panel) of the polyacrylamide-bound Mn^{2+}-Phos-tag^{TM} ligand (25 μM) was performed (9 μg of whole cell [whole cell lysis buffer] or cytoplasmic fractions and 2 μg of nuclear fraction, based on the hnRNP K content for clarity of display) followed by immunoblotting with an anti-hnRNP K/J antibody. Each fraction was obtained from J774.1 cells that were either left untreated (LPS−) or incubated in the presence of 100 ng/mL LPS for 4 h (LPS+) (C and D). The panel D shows endogenous hnRNP K following λ-PPase treatment for dephosphorylation. The incubation time for each reaction was 0, 10, or 60 min. The white and black arrows indicated the bands that disappeared and remained after dephosphorylation, respectively. "Whole", "Cyto", and "Nuc" indicate whole cell extracts, cytoplasmic fraction, and nuclear fraction, respectively.

3.2.1 Multiple forms of hnRNP K generated by alternative splicing

In the UniProtKB/SwissProt database (http://www.uniprot.org/), hnRNP K is described as having three isoforms that are derived by alternative splicing of the *Hnrnpk* gene (isoforms 1, 2, and 3 in Fig. 3B). At the protein level, isoforms 1 and 2 have a 99.8% overall coding sequence identity with small differences in the C-terminus (isoform 1, a basic C-terminal end, ^{459}SGKFF^{463}; isoform 2, an acidic C-terminal end, ^{459}ADVEGF^{464}). On the other hand, the C-terminal sequence of isoform 3 is identical to that of isoform 1, but an internal single exon encoding the KH1 and KH2 domains is absent. Thus, by comparing the MS results of the peptides (I and/or C region) containing the internal skipping exon and/or C-terminal exons, respectively, of hnRNP K, we could specify which alternative forms accumulated in J774.1 cells (Fig. 3C and Supporting Information Table S1). The maximum peptide ion scores of the detected peptides in these regions for each form are shown in Table 1. The MS/MS spectra of these peptides enabled us to precisely assign the exact peptide (Fig. 3D). Although several target peptides in some of the 20 spots in Fig. 3A were not detected by MS (Table 1), the relative positions of these spots on the phosphate-affinity 2-DE were also taken into account when defining the products generated by alternative splicing. Consequently, each of the alternatively spliced isoforms was separately grouped into four areas that are designated with dotted lines on the phosphate-affinity 2-DE map in Fig. 3A (Table 1). These results indicated that there was an unidentified isoform, designated isoform 4, that had the isoform 2-type C-region but lacked the same I region as in isoform 3 (Fig. 3B). To confirm the presence of isoform 4 in J774.1 cells, we cloned the coding sequences of hnRNP K by RT-PCR and examined the number of alternatively spliced forms. Consistent with the MS results, the subcloned RT-PCR products confirmed that there were four different products generated by alternative splicing (data not shown). Therefore, these results revealed that the presence of four alternatively spliced forms of hnRNP K was one reason why multiple protein spots were detected by phosphate-affinity 2-DE.

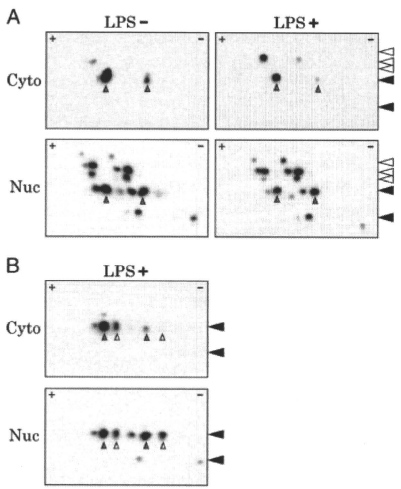

Figure 2. Multiple forms of endogenous hnRNP K on 2-DE gel images in the presence of the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand (25 μM). Each fraction was separated by mini 2-DE with narrow-range IPG gel (pH 4.7–5.9) in the first dimension and 7.5% w/v polyacrylamide gel in the presence of the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand (25 μM) in the second dimension followed by immunoblotting with an anti-hnRNP K/J antibody. Each fraction was obtained from J774.1 cells that were either left untreated (LPS–) or treated with 100 ng/mL LPS for 4 h (LPS+). The fractions prepared from LPS-stimulated cells following λ-PPase-mediated dephosphorylation are shown in the (B). Black and white arrows indicate the vertical positions of the horizontally contiguous spots corresponding to those disappeared and remained, respectively, after dephosphorylation on the phosphate-affinity SDS-PAGE gel images shown in Fig. 1D. In (A and B), gray and white arrows indicate the positions that correspond to spot 10 or 12 and spot 23 or 26 (basic spot) in Fig. 4A, respectively. "Cyto" and "Nuc" indicate the cytoplasmic and nuclear fractions, respectively.

3.2.2 Multiple forms of hnRNP K generated by phosphorylation

To determine how the degree of phosphorylation relates to the multiple forms of endogenous hnRNP K, we next analyzed the phosphorylation states of each spot. HnRNP K is reported to be tyrosine phosphorylated by Src-family kinases [13, 19]. However, hnRNP K phosphorylation at Tyr residues in LPS-stimulated J774.1 cells was not detected when hnRNP K was immunoprecipitated, analyzed by phosphate-affinity SDS-PAGE, and then examined by immunoblotting analysis with an anti-phosphotyrosine antibody (data not shown). Thus, Tyr phosphorylation of hnRNP K was not prominent in J774.1 cells. On the other hand, hnRNP K is also known to be phosphorylated at multiple Ser/Thr residues by kinase(s) (Fig. 3B) [13]. Therefore, we focused analysis of phosphorylation state on Ser/Thr residues in hnRNP K by MS. In particular, the phosphorylation states of isoforms 1 and 2 were investigated in detail because of their abundance. In the present study, the MS analysis characterized several peptides that contained Ser or Thr residues and were shown to be phosphorylated in UniProtKB and/or PhosphoSite (Fig. 3B and Supporting Information Table S2). For each form, the maximum peptide ion score of the detected peptides that contain Ser or Thr residues and were shown to be phosphorylated in UniProtKB and/or PhosphoSite are shown in Table 1. The MS/MS spectra of these peptides enabled us to precisely assign the exact peptide. Consequently, the MS analysis indicated that spots 3, 4, and 6 and spots 7 and 8 were phosphorylated at Ser116 and Ser284, respectively, while spot 2 was phosphorylated at both Ser116 and Ser284 (Table 1 and Fig. 3A). Apparently, forms with the same migration rates in the vertical axis on phosphate-affinity SDS-PAGE were phosphorylated at the same residues. If this rule is generally correct, spots 1 and 5 are expected to be phosphorylated at Ser116/Ser284 and Ser116, respectively, although the MS analysis did not identify these phosphorylated peptide(s). On the other hand, the phosphorylated peptides for spots 9, 10, 11, and 12 and all of the forms derived from either isoform 3 or 4 were very scarce, and thus were not detected (Table 1).

Because a peptide containing Ser353, which is predicted to be phosphorylated [17, 18], was not detected by MS, we analyzed a series of mutants to determine whether hnRNP K is phosphorylated at Ser353 or not. In this experiment, this Ser residue was replaced with Ala by mutating the DNA of an hnRNP K expression clone. Substituting Ser with Ala is expected to completely block phosphorylation at this amino acid residue but minimally affect the electrophoretic mobility of hnRNP K in phosphate-affinity 2-DE. Each mutant is designated based on the amino acid substitution and the mutated residue number (for example, a mutant where Ser116 was converted into Ala is designated S116A). Thus, a comparison of the electrophoretic mobilities of these mutant and endogenous hnRNP K proteins would clarify whether this specific Ser residue is phosphorylated *in vivo*. To produce these wild-type and mutant proteins, we used a wheat germ cell-free protein synthesis system. This wheat germ system is ideal because it is derived from an eukaryotic organism and efficiently produces recombinant proteins but does not contain detectable levels of hnRNP K, as determined by probing the extract with anti-mammalian hnRNP K/J antibody. However, before using this system, we first examined the PTMs of hnRNP K that are produced in the wheat germ system. Interestingly, the wheat germ system yielded hnRNP K that showed multiple spots on phosphate-affinity 2-DE, and most of these spots were in identical positions to those of mouse endogenous hnRNP K. The major differences were that spots 3, 5, 9, and 11 disappeared while spots 21, 22, 23, 24, 25, and 26 newly appeared (Figs. 3A and 4A). In addition, because a

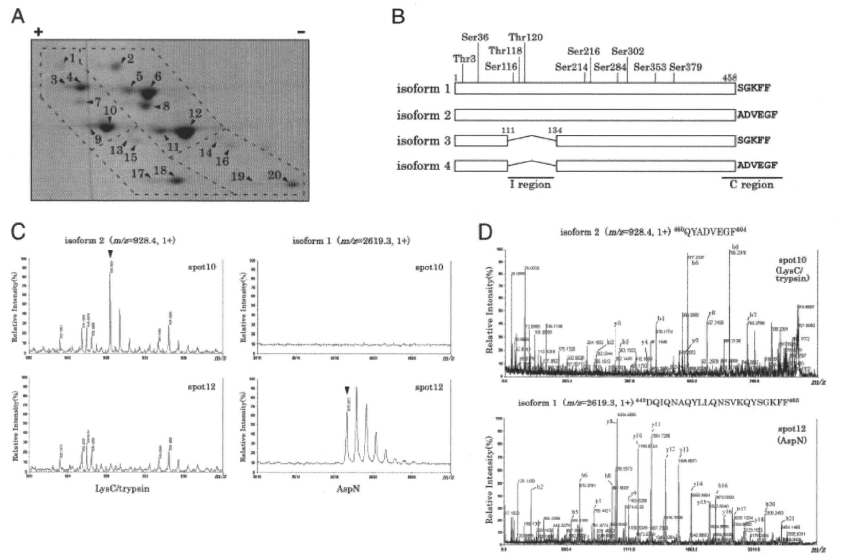

Figure 3. The identification of multiple forms of endogenous hnRNP K generated by alternative splicing. (A) Expanded SYPRO Ruby-stained 2-DE gel image of endogenous hnRNP K isolated from nuclear fractions of J774.1 cells that were incubated with 100 ng/mL LPS for 4 h. The extracts were immunoprecipitated with an anti-hnRNP K/J antibody. The isolated proteins were separated by mini 2-DE in the presence of the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand (25 μM) as described in Fig. 2, and subsequently detected by SYPRO Ruby staining. The four areas surrounded by a dotted square that corresponds to each isoform (Fig. 5A). (B) A schematic map of the isoforms generated by alternative splicing of the single hnRNP K gene. I and C regions indicate the peptide regions coded by the internal skipping exon and C-terminal exons of hnRNP K, respectively. (C) MS data from LysC/trypsin digests (left panel) and AspN digests (right panel) of spots 10 and 12. Mass spectra of *m/z* 928.4 or *m/z* 2619.3 were only detected for isoforms 2 or 1, respectively. (D) Tandem mass spectra of *m/z* 928.4 or *m/z* 2619.3 signals. Each sequence corresponding to the C-terminal peptide of hnRNP K isoform 2 (upper panel) or 1 (bottom panel) was identified.

mutant hnRNP K with S116A/S284A/S353A did not show any phosphorylated forms on the phosphate-affinity 2-DE map, Ser116, Ser284, and Ser353 were considered to be main phosphoacceptor sites of hnRNP K produced in the wheat germ system as in the case of mammalian hnRNP K.

Mutant forms of hnRNP K isoform 1 or 2 (carrying mutations S116A, S284A, and S116A/S284A) were produced and analyzed using the wheat germ system. The phosphate-affinity 2-DE gel image indicated that spots 1, 2, 4, 6, 7, and 8 disappeared compared with that of wild-type hnRNP K isoforms 1 or 2. These findings were consistent with our MS analysis, which showed that the proteins in these spots were phosphorylated at Ser116 and/or Ser284. Thus, this analysis indicated that the wheat germ system could serve as a reliable method to indirectly determine the phosphorylation states of hnRNP K. Thus, the hnRNP K mutants (S353A) of isoforms 1 and 2 were examined; the products did not give spots 21 and 22 (isoform 1) or spots 24 and 25 (isoform 2), whereas spots 1, 2, 4, 6, 7, and 8 remained unchanged (Fig. 4A). These results suggested that spots 21/24 or spots 22/25 contained an hnRNP K form that was phosphorylated at Ser353 or Ser284/Ser353, respectively, while the other spots were unlikely to include hnRNP K forms that were phosphorylated at Ser353.

3.2.3 Multiple forms of hnRNP K generated by PTMs other than phosphorylation

Many of the alternative forms of endogenous hnRNP K that are generated by phosphorylation disappeared upon dephosphorylation with λ-PPase treatment. However, we observed that two new basic spots emerged upon λ-PPase treatment (Fig. 2B). When the wild-type hnRNP K was synthesized in the wheat germ cell-free system and analyzed without dephosphorylation, spots (23 and 26) that corresponded to the basic

Table 1. The identification of the multiple forms of hnRNP K produced by alternative splicing and phosphorylation

Spot no.	The deduced modification site[a]										Variant region generated by alternative splicing[b]			
	Ac-M1/ T3	S36	S116/–T118/ T120	*p*-S116/–T118/ T120	S214/ S216	S284	*p*-S284	dm-R296/dm-R299/S302	dm-R303	S379	I1	I2	C1	C2
1	95 (42)				63 (39)					143 (42)				
2	133 (42)	61 (35)		44 (42)	74 (37)		40 (33)			142 (42)	44 (42)		81 (38)	
3	99 (42)			74 (42)	95 (38)	41 (36)		61 (41)	61 (41)	166 (42)	74 (42)			54 (37)
4	130 (43)	69 (35)		82 (42)	97 (39)	40 (35)		69 (42)	68 (42)	147 (42)	82 (42)			63 (31)
5	92 (39)	82 (35)			96 (37)	37 (36)				130 (42)			117 (38)	
6	154 (39)	78 (35)		121 (42)	106 (34)			53 (41)		168 (42)	121 (42)		131 (38)	
7	110 (41)				74 (38)		40 (33)			133 (42)				
8	111 (41)	68 (35)			78 (38)		41 (32)	44 (41)		111 (42)			55 (38)	
9	81 (42)	56 (37)	63 (42)		95 (40)		39 (35)	44 (43)	44 (43)	157 (42)	63 (42)			59 (30)
10	136 (42)	56 (35)	136 (42)		103 (40)		41 (30)	61 (43)	61 (43)	166 (42)	136 (42)			99 (36)
11	66 (41)	84 (35)			83 (40)					146 (42)			92 (38)	
12	153 (39)	86 (35)	118 (42)		107 (34)	36 (35)		56 (41)		151 (42)	118 (42)		126 (38)	
13	116 (41)	66 (35)			62 (35)					110 (42)		58 (41)		
14	121 (42)	78 (35)			54 (35)					120 (42)				
15	119 (43)	74 (35)			44 (40)			59 (42)		94 (42)		59 (41)		110 (37)
16	128 (43)	66 (35)			51 (39)					99 (42)		62 (41)		
17		87 (35)			78 (35)			53 (42)		152 (41)		66 (41)		
18	141 (43)	86 (35)			101 (35)			75 (42)		150 (42)		65 (41)		44 (36)
19								50 (42)		110 (42)		63 (41)		
20	152 (38)	88 (35)			84 (38)			51 (42)		153 (42)		66 (41)		

The number indicates the maximum peptide match score of the detected peptide including each residue and the number given in parentheses indicates their individual peptide score identity threshold.

a) These deduced modification sites are referred to PhosphoSite and UniProtKB/SwissProt database and the detected peptide list is shown in Supporting Information Table S2. The representative MS/MS spectra of the peptide including PTM sites are indicated in Supporting Information Fig. S1. Ac-, *p*-, and dm- represent *N*-terminal acetylation, phosphorylation, and dimethylation, respectively.

b) The detected peptide list of these variant regions is shown in Supporting Information Table S1.

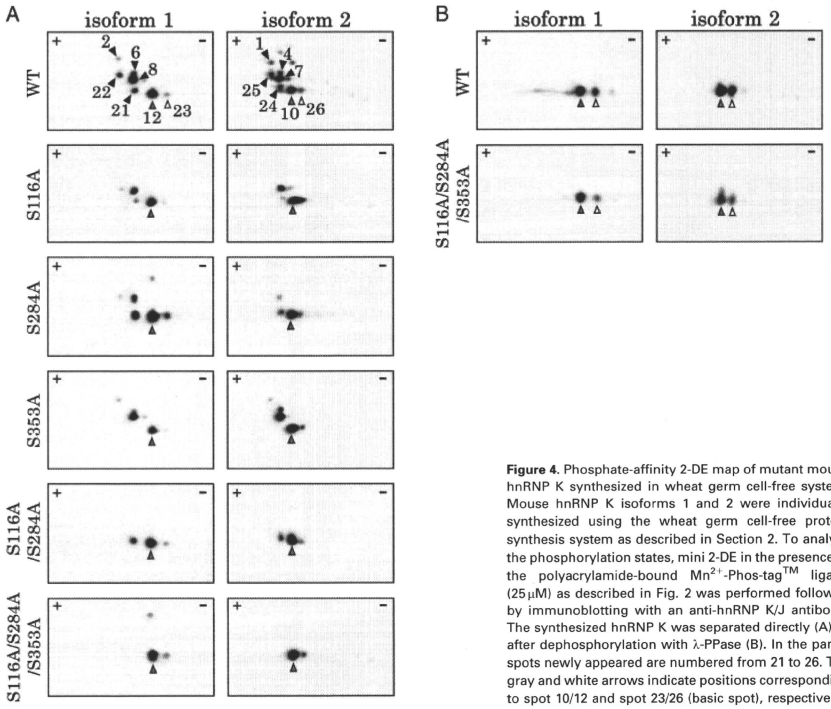

Figure 4. Phosphate-affinity 2-DE map of mutant mouse hnRNP K synthesized in wheat germ cell-free system. Mouse hnRNP K isoforms 1 and 2 were individually synthesized using the wheat germ cell-free protein synthesis system as described in Section 2. To analyze the phosphorylation states, mini 2-DE in the presence of the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand (25 μM) as described in Fig. 2 was performed followed by immunoblotting with an anti-hnRNP K/J antibody. The synthesized hnRNP K was separated directly (A) or after dephosphorylation with λ-PPase (B). In the panel, spots newly appeared are numbered from 21 to 26. The gray and white arrows indicate positions corresponding to spot 10/12 and spot 23/26 (basic spot), respectively.

spots of dephosphorylated hnRNP K were detected. After λ-PPase-mediated dephosphorylation, these basic spots of S116A/S284A/S353A in isoforms 1 and 2 were detected at a lower level than those of the dephosphorylated wild-type protein because there was a sharp decrease in the phosphorylation state of this mutant (Fig. 4B). These results indicated that these basic spots were not an artifact associated with λ-PPase dephosphorylation and that hnRNP K in these spots were modified by some PTMs other than phosphorylation. Because the expression levels of the detected phosphorylated forms of hnRNP K were not consistent with those of the basic spots after dephosphorylation, phosphorylation and the other PTMs could co-occur on the same hnRNP K molecule.

HnRNP K is reportedly dimethylated by protein arginine methyltransferase 1 [21]. Consistent with this report, we observed MS signals corresponding to peptides containing dimethylated residues at Arg296, Arg299, and Arg303, in LysC- and/or trypsin-digests derived from spots 3, 4, 6, 8, 9, 10, 15, 17, 18, 19, and 20 (Table 1). MS analysis of trypsin- or AspN-digests derived from 18 of 20 spots, which excluded spots 17 and 19, showed that N-terminal acetylation occurred at Met1 (Table 1). However, these PTMs were observed in almost all the horizontally contiguous spots with a similar migration pattern along the vertical axis on phosphate-affinity 2-DE. The cooperative PTMs analysis of immunoprecipitated hnRNP K using ProteinPilot version 3.0 (Applied Biosystems) failed to detect more PTMs other than them. Thus, the distinct PTMs causing the p*I* gel shifts of hnRNP K could not be exclusively determined yet.

3.2.4 Subcellular localization of the multiple hnRNP K forms

Based on the results described above, the structures of the multiple forms of endogenous hnRNP K detected on phosphate-affinity 2-DE gel images are schematically summarized in Fig. 5. The nucleus contained four hnRNP K isoforms that

were generated by alternative splicing, and isoforms 1 and 2 were the major alternatively spliced forms (Fig. 3A). Interestingly, isoform 2 preferentially accumulated in the cytoplasm compared to isoform 1, whereas there were equal amounts of these isoforms in the nucleus. Isoforms 3 and 4, which lack a single internal exon that is present in isoforms 1 and 2, were not detected in the cytoplasm (Fig. 2A). For isoforms 1 and 2, multiple forms that were phosphorylated at Ser116, Ser284, or Ser116/Ser284 were predominantly expressed. In the nucleus, the Ser116/Ser284 or Ser116 phosphorylated forms increased in LPS-stimulated cells, while the Ser284 phosphorylated forms and non-phosphorylated forms decreased. On the other hand, the cytoplasm contained higher levels of the Ser116 phosphorylated forms, especially isoform 2, following LPS stimulation (Fig. 2A). Therefore, these results indicated that multiple forms of hnRNP K were produced mainly by alternative splicing and phosphorylation of Ser116 and/or Ser284, and these forms showed different responses and subcellular localizations after LPS stimulation. The Ser353 phosphorylated form was minor at least in J774.1 cells.

3.3 Modulation of the contents of the multiple hnRNP K forms upon serum stimulation

Phosphate-affinity 2-DE analysis of hnRNP K showed that LPS stimulation modified the multiple forms of hnRNP K

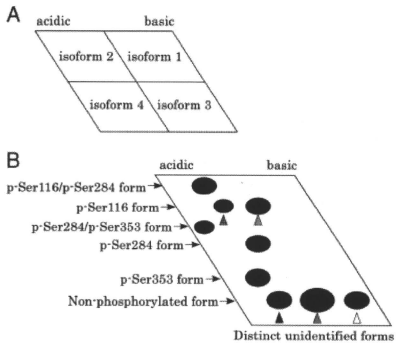

Figure 5. Schematic illustration of the phosphate-affinity 2-DE map of hnRNP K. The isoform-specific regions generated by alternative splicing (Fig. 3A) are illustrated in the (A). The relative detected positions of multiple phosphorylated forms of hnRNP K on each region of isoforms 1 or 2 are shown in (B) based on the data using J774.1 cells and the wheat germ cell-free protein synthesis system. p-Ser represents phosphorylated Ser. The vertical position indicates the distinct phosphorylation form while non-phosphorylated and pSer116-phosphorylated forms appeared multiple horizontal spots (highlighted by arrowheads) probably because of unidentified PTMs.

that accumulated in the nucleus and the cytoplasm. Therefore, we examined the changes in the PTMs of hnRNP K using a phosphate-affinity 2-DE map in order to understand the functional significance of these PTMs. In a previous report, serum stimulation of HeLa cells resulted in the phosphorylation and the cytoplasmic accumulation of hnRNP K [17]. ERK, which is activated after serum stimulation, was described to be responsible for phosphorylation at Ser284 and Ser353 of hnRNP K from the data using mutant hnRNP K [18]. However, it has not been demonstrated how phosphorylation states of endogenous hnRNP K are modulated in HeLa cells after serum stimulation by a direct protein–chemical method. We thus characterized the PTM patterns of endogenous hnRNP K in HeLa cells before and after serum stimulation using the phosphate-affinity 2-DE map obtained in this study. Because the coding sequence of hnRNP K is completely conserved between humans and mice, the positions of multiple spots of endogenous hnRNP K in HeLa cells on the phosphate-affinity 2-DE map were perfectly superimposable with those in J774.1 cells (data not shown). This enabled us to specify the phosphorylation state of hnRNP K in HeLa cells by simply comparing the spot positions on the phosphate-affinity 2-DE map with those from J774.1 cells. The cytoplasm of HeLa cells contains higher levels of isoform 1 than J774.1 cells. Conventional SDS-PAGE immunoblotting analyses revealed that the levels of endogenous hnRNP K in the cytoplasm increased approximately twofold after serum stimulation as described in the previous report (Fig. 6B) [17]. Phosphate-affinity 1-DE and 2-DE also showed that the Ser116 phosphorylated forms of endogenous hnRNP K, especially for isoform 2, increased in the nucleus and the cytoplasm after serum stimulation (Fig. 6C). However, the Ser284 and/or Ser353 phosphorylated forms remained constant in the nucleus and cytoplasm of HeLa cells while ERK was activated after serum stimulation (Fig. 6A). These results suggested that hnRNP K in HeLa cells was preferentially phosphorylated at Ser116 compared to Ser284 or Ser353 after serum stimulation, which was similar to the changes in J774.1 cells after LPS stimulation.

4 Discussion

Protein activity is not solely controlled by the rates of protein biosynthesis and degradation. PTMs functionally modulate proteins, and specific and selective phosphorylation is a well-known example of PTMs [1, 2]. Thus, one of the major focuses in proteomics is to conduct a comprehensive analysis of the phosphorylated forms of proteins. In the present study, we structurally characterized multiple forms of endogenous hnRNP K using a combination of phosphate-affinity 2-DE and MS analysis. Based on the phosphate-affinity 2-DE map of hnRNP K generated in this study, we could delineate the relationship between phosphorylation and the differential subcellular localization of multiple

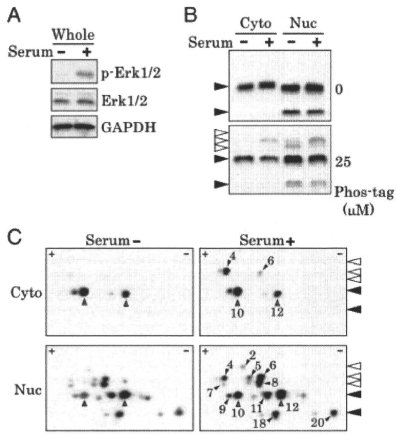

Figure 6. The modulation of multiple forms of endogenous HnRNP K after serum stimulation of serum-starved HeLa cells. Each fraction was obtained from serum-starved HeLa cells that were either left untreated (serum–) or stimulated with serum for 30 min (serum+). "Whole", "Cyto", and "Nuc" indicate whole cell extracts, cytoplasmic fraction, and nuclear fraction, respectively. (A) The whole cell extracts [SDS-PAGE sample buffer] were immunoblotted for the total expression levels and phosphorylation states of ERK1/2. (B) To examine the phosphorylation states of endogenous hnRNP K, 7.5% w/v SDS-PAGE in the absence (upper panel) or presence (bottom panel) of the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand (25 μM) was performed (9 μg of cytoplasmic fraction and 2 μg of nuclear fraction, based on hnRNP K content for clarity of display) followed by immunoblotting with an anti-hnRNP K/J antibody. (C) Each fraction was separated by mini 2-DE in the presence of the polyacrylamide-bound Mn^{2+}-Phos-tag™ ligand (25 μM) as described in Fig. 2 and subsequently detected by immunoblotting with an anti-hnRNP K/J antibody. The black and white arrows indicate the trains corresponding to bands that disappeared or remained, respectively, on the phosphate-affinity SDS-PAGE gel images shown in Fig. 1D. In the panel, the gray arrows indicate spots corresponding to spots 10 or 12.

phosphorylated/alternatively spliced forms of hnRNP K and determine how these characteristics were modulated upon external stimulation. The combination of phosphate-affinity 2-DE and MS analysis provided significant and abundant information about multiple forms of hnRNP K, which had been previously only indirectly determined either by extensive LC-MS techniques or mutational analyses.

HnRNP K has diverse roles in regulating gene expression at a variety of transcriptional and translational levels. Many of these processes are regulated by hnRNP K phosphorylation that is induced by external stimulation. Using a point mutational analysis, Habelhah *et al.* [17, 18] showed that ERK-mediated phosphorylation of Ser284 and Ser353 contributes to nuclear export of hnRNP K and a concomitant inhibition of translation, while JNK-mediated phosphorylation of Ser216 and Ser353 increases the transcriptional effects of hnRNP K. Although these mutational analyses are powerful approaches that determine whether PTMs at specific site(s) affect protein subcellular localization and/or function or not, they may not necessarily reflect PTMs occurring on endogenous proteins in response to external stimulation. For example, we found that transient heterologous expression of hnRNP K itself enhanced its phosphorylation at Ser116 (data not shown). Thus, the results obtained by using transiently overexpressed mutants must be carefully interpreted. As an alternative and direct approach, we performed a direct protein–chemical characterization of PTMs of hnRNP K using phosphate-affinity electrophoresis and revealed that phosphorylation at Ser116 and/or Ser284 of hnRNP K was modulated by external stimulation. In particular, phosphorylation of Ser116 was enhanced in the cytoplasm and the nucleus, suggesting that phosphorylation at Ser116 of endogenous hnRNP K might be essential for the intracellular function of hnRNP K after external stimulation. In this regard, it is interesting to explore the function of isoforms 3 and 4 of hnRNP K because they do not carry a region including Ser116 and are specifically localized in the nucleus as shown in this study for the first time. Phosphorylation of Ser284 might also play an important role in the nucleus because these forms specifically accumulate in the nuclear fraction independent of external stimulation. However, these phosphorylation sites were not present in the nuclear localization signal and K nuclear shuttling domains or three KH domains.

Phosphate-affinity 2-DE technology is a new technology that was successfully used in this study to generate a structural variation map of endogenous hnRNP K. Interestingly, phosphate-affinity 2-DE provides significantly enhanced resolution of phosphorylated proteins because the electrophoretic mobility of proteins depends not only on the presence of a phosphate moiety on an amino acid residue but also on the sequence of the phosphorylated amino acid residue. This enhanced resolving power enabled us to identify additional structural variations due to alternative splicing or unknown PTMs and will significantly advance our ability to analyze PTMs in structural detail. In fact, subtle differences in phosphorylation patterns of isoforms 1 and 2 of hnRNP K elicited by external stimulation (*e.g.* with LPS or serum) could not be detected without the phosphate-affinity 2-DE mapping (Figs. 2A and 6C).

The next step is to link structural variations with protein function. For example, we found that each structural variant of hnRNP K preferentially localized to specific intracellular sites and differentially responded to external stimulation with LPS or serum. Because hnRNP K has a wide range of biological functions in mammalian cells, additional data on the structural variations of hnRNP K under various biolo-

gical conditions will help determine the relationship between PTMs/alternative splicing and the biological functions of hnRNP K. To this end, the phosphate-affinity 2-DE map of hnRNP K generated in this study will provide critical information.

This work was partially supported by Special Coordination Funds for Promoting Science and Technology "Creation of Innovation Centers for Advanced Interdisciplinary Research Areas".

The authors have declared no conflict of interest.

5 References

[1] Mann, M., Jensen, O. N., Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 2003, *21*, 255–261.

[2] Jensen, O. N., Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* 2004, *8*, 33–41.

[3] Aebersold, R., Mann, M., Mass spectrometry-based proteomics. *Nature* 2003, *422*, 198–207.

[4] Schmelzle, K., White, F. M., Phosphoproteomic approaches to elucidate cellular signaling networks. *Curr. Opin. Biotechnol.* 2006, *17*, 406–414.

[5] Görg, A., Weiss, W., Dunn, M. J., Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 2004, *4*, 3665–3685.

[6] Morandell, S., Stasyk, T., Grosstessner-Hain, K., Roitinger, E., *et al.*, Phosphoproteomics strategies for the functional analysis of signal transduction. *Proteomics* 2006, *6*, 4047–4056.

[7] Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K., Koike, T., Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* 2006, *5*, 749–757.

[8] Kinoshita-Kikuta, E., Aoki, Y., Kinoshita, E., Koike, T., Label-free kinase profiling using phosphate affinity polyacrylamide gel electrophoresis. *Mol. Cell. Proteomics* 2007, *6*, 356–366.

[9] Kinoshita, E., Kinoshita-Kikuta, E., Matsubara, M., Aoki, Y. *et al.*, Two-dimensional phosphate-affinity gel electrophoresis for the analysis of phosphoprotein isotypes. *Electrophoresis* 2009, *30*, 550–559.

[10] Michael, W. M., Eder, P. S., Dreyfuss, G., The K nuclear shuttling domain: a novel signal for nuclear import and nuclear export in the hnRNP K protein. *EMBO J.* 1997, *16*, 3587–3598.

[11] Dzwonek, A., Mikula, M., Ostrowski, J., The diverse involvement of heterogeneous nuclear ribonucleoprotein K in mitochondrial response to insulin. *FEBS Lett.* 2006, *580*, 1839–1845.

[12] Bomsztyk, K., Van Seuningen, I., Suzuki, H., Denisenko, O., Ostrowski, J., Diverse molecular interactions of the hnRNP K protein. *FEBS Lett.* 1997, *403*, 113–115.

[13] Bomsztyk, K., Denisenko, O., Ostrowski, J., hnRNP K: one protein multiple processes. *BioEssays* 2004, *26*, 629–638.

[14] Makeyev, A. V., Liebhaber, S. A., The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. *RNA* 2002, *8*, 265–278.

[15] Van Seuningen, I., Ostrowski, J., Bustelo, X. R., Sleath, P. R., Bomsztyk, K., The K protein domain that recruits the interleukin 1-responsive K protein kinase lies adjacent to a cluster of c-Src and Vav SH3-binding sites. Implications that K protein acts as a docking platform. *J. Biol. Chem.* 1995, *270*, 26976–26985.

[16] Schullery, D. S., Ostrowski, J., Denisenko, O. N., Stempka, L. *et al.*, Regulated interaction of protein kinase Cdelta with the heterogeneous nuclear ribonucleoprotein K protein. *J. Biol. Chem.* 1999, *274*, 15101–15109.

[17] Habelhah, H., Shah, K., Huang, L., Ostareck-Lederer, A. *et al.*, ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. *Nat. Cell Biol.* 2001, *3*, 325–330.

[18] Habelhah, H., Shah, K., Huang, L., Burlingame, A. L. *et al.*, Identification of new JNK substrate using ATP pocket mutant JNK and a corresponding ATP analogue. *J. Biol. Chem.* 2001, *276*, 18090–18095.

[19] Ostareck-Lederer, A., Ostareck, D. H., Cans, C., Neubauer, G., *et al.*, c-Src-mediated phosphorylation of hnRNP K drives translational activation of specifically silenced mRNAs. *Mol. Cell Biol.* 2002, *22*, 4535–4543.

[20] Li, T., Evdokimov, E., Shen, R. F., Chao, C. C. *et al.*, Sumoylation of heterogeneous nuclear ribonucleoproteins, zinc finger proteins, and nuclear pore complex proteins: a proteomic analysis. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 8551–8556.

[21] Ostareck-Lederer, A., Ostareck, D. H., Rucknagel, K. P., Schierhorn, A.*et al.*, Asymmetric arginine dimethylation of heterogeneous nuclear ribonucleoprotein K by protein-arginine methyltransferase 1 inhibits its interaction with c-Src. *J. Biol. Chem.* 2006, *281*, 11115–11125.

[22] Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 1976, *72*, 248–254.

[23] Hijikata, A., Kitamura, H., Kimura, Y. *et al.*, Construction of an open-access database that integrates cross-reference information from the transcriptome and proteome of immune cells. *Bioinfomatics* 2007, *23*, 2934–2941.

Survival Prediction for Pancreatic Cancer Patients Receiving Gemcitabine Treatment*[S]

Junichi Matsubara‡§¶, Masaya Ono‡, Kazufumi Honda‡, Ayako Negishi‡, Hideki Ueno‖, Takuji Okusaka‖, Junji Furuse**, Koh Furuta‡‡, Emiko Sugiyama§§, Yoshiro Saito§§, Nahoko Kaniwa§§, Junichi Sawada§§, Ayako Shoji¶¶, Tomohiro Sakuma¶¶, Tsutomu Chiba§, Nagahiro Saijo‖‖, Setsuo Hirohashi‡, and Tesshi Yamada‡

Although gemcitabine monotherapy is the standard treatment for advanced pancreatic cancer, patient outcome varies significantly, and a considerable number do not benefit adequately. We therefore searched for new biomarkers predictive of overall patient survival. Using LC-MS, we compared the base-line plasma proteome between 29 representative patients with advanced pancreatic cancer who died within 100 days and 31 patients who survived for more than 400 days after receiving at least two cycles of the same gemcitabine monotherapy. Identified biomarker candidates were then challenged in a larger cohort of 304 patients treated with the same protocol using reverse-phase protein microarray. Among a total of 45,277 peptide peaks, we identified 637 peaks whose intensities differed significantly between the two groups ($p < 0.001$, Welch's t test). Two MS peaks with the highest statistical significance ($p = 2.6 \times 10^{-4}$ and $p = 5.0 \times 10^{-4}$) were revealed to be derived from α_1-antitrypsin and α_1-antichymotrypsin, respectively. The levels of α_1-antitrypsin ($p = 8.9 \times 10^{-8}$) and α_1-antichymotrypsin ($p = 0.001$) were significantly correlated with the overall survival of the 304 patients. We selected α_1-antitrypsin ($p = 0.0001$), leukocyte count ($p = 0.066$), alkaline phosphatase ($p = 8.3 \times 10^{-8}$), and performance status ($p = 0.003$) using multivariate Cox regression analysis and constructed a scoring system (nomogram) that was able to identify a group of high risk patients having a short median survival time of 150 days (95% confidence interval, 123–187 days; $p = 2.0 \times 10^{-15}$, log rank test). The accuracy of this model for prognostication was internally validated and showed good calibration and discrimination with a bootstrap-corrected concordance index of 0.672. In conclusion, an increased level of α_1-antitrypsin is a biomarker that predicts short overall survival of patients with advanced pancreatic cancer receiving gemcitabine monotherapy. Although an external validation study will be necessary, the current model may be useful for identifying patients unsuitable for the standardized therapy. ***Molecular & Cellular Proteomics 9:695–704, 2010.***

Invasive ductal adenocarcinoma of the pancreas is one of the most aggressive and lethal malignancies (1). It is the fifth leading cause of cancer-related death in Japan and the fourth in the United States, accounting for an estimated >23,000 deaths per year in Japan and >33,000 in the United States (2, 3). Because the majority of patients have distant metastases even at their first presentation (4, 5), the main therapeutic modality for pancreatic cancer is systemic chemotherapy, and gemcitabine monotherapy is the current standard (6). Gemcitabine treatment has significantly improved the median survival time of patients with advanced pancreatic cancer (7). However, the outcome of the treatment varies significantly among individuals, and a considerable portion of patients do not appear to benefit significantly from it. It therefore seems necessary to assess the efficacy and adverse effects of the drug before administration and tailor the treatment accordingly for each person.

We previously identified a predictive biomarker for hematologic toxicity, which is one of the most frequent and potentially life-threatening adverse effects associated with gemcitabine monotherapy (8). As a next step, we performed a large scale proteome analysis in this study to identify biomarkers predictive of patient survival after gemcitabine monotherapy. Several factors and their combinations have been reported to correlate significantly with outcome in patients with advanced pancreatic cancer receiving gemcitabine, such as performance status, metastases, serum albumin, alkaline phosphatase, and peripheral leukocyte count (9–11). Unfortunately, however, the accuracy of survival prediction based on these conventional prognostic factors seems unsatisfactory (9).

In recent years, there has been considerable interest in applying advanced proteomics technologies to the discovery of predictive biomarkers (12, 13). We and others have

From the ‡Chemotherapy Division, National Cancer Center Research Institute, Tokyo 104-0045, ‖Hepatobiliary and Pancreatic Oncology Division and ‡‡Clinical Laboratory Division, National Cancer Center Hospital, Tokyo 104-0045, §§Project Team for Pharmacogenetics, National Institute of Health Sciences, Tokyo 158-8501, ¶¶BioBusiness Group, Mitsui Knowledge Industry, Tokyo 105-6215, **Hepatobiliary and Pancreatic Oncology Division, ‖‖National Cancer Center Hospital East, Kashiwa 277-8577, and §Department of Gastroenterology and Hepatology, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan

Received, May 12, 2009, and in revised form, January 6, 2010
Published, MCP Papers in Press, January 8, 2010, DOI 10.1074/mcp.M900234-MCP200

TABLE I

Clinical and laboratory data of patients with short term or long term survival

Survival time was calculated from the date of starting gemcitabine therapy until the date of death from cancer. Wilcoxon test was applied to assess differences in values. 5-FU, 5-fluorouracil; LAPC, locally advanced pancreatic cancer; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; C_{max}, peak concentration; AUC, area under the curve.

	Short term survivor (<100 days)	Long term survivor (>400 days)	*p*
No. of patients	29	31	
Sex (no. of patients)			0.361[a]
Male	21	19	
Female	8	12	
Age, mean (S.D.) (years)	63 (7)	67 (8)	0.123
ECOG performance status (no. of patients)			0.008[a]
0	8	20	
1	18	11	
2	3	0	
Body surface area, mean (S.D.) (m^2)	1.59 (0.17)	1.54 (0.15)	0.333
Prior therapy			0.438[a]
None	27	27	
Chemoradiotherapy using 5-FU for LAPC	2	4	
Clinical stage[b]			0.697[a]
IVa	2	3	
IVb	27	28	
Subsequent line chemotherapy after gemcitabine			0.045[a]
None	29	27	
Yes	0	4	
Leukocytes, mean (S.D.) ($\times 10^3/mm^3$)	7.6 (3.6)	5.2 (1.3)	0.002
Platelets, mean (S.D.) ($\times 10^4/mm^3$)	24.5 (7.6)	20.2 (4.6)	0.020
Hemoglobin, mean (S.D.) (g/dl)	11.7 (1.6)	11.7 (1.5)	0.491
Albumin, mean (S.D.) (g/dl)	3.4 (0.4)	3.7 (0.3)	0.014
Creatinine, mean (S.D.) (mg/dl)	0.70 (0.23)	0.68 (0.23)	0.726
AST, mean (S.D.) (IU/liter)	40 (25)	26 (15)	0.010
ALT, mean (S.D.) (IU/liter)	51 (44)	27 (19)	0.037
ALP, mean (S.D.) (units/liter)	728 (632)	337 (160)	0.026
Pharmacokinetic parameters of gemcitabine			
C_{max}, mean (S.D.) (μg/ml)	24.02 (7.52)	24.91 (6.22)	0.610
AUC, mean (S.D.) (h·μg/ml)	10.24 (2.83)	10.75 (2.32)	0.270
α_1-Antitrypsin,[c] mean (S.D.)	64.6 (66.8)	16.9 (7.9)	0.0003
α_1-Antichymotrypsin,[c] mean (S.D.)	706.4 (416.0)	389.0 (216.5)	0.0005
Tumor response[d]			<0.0001[a]
Complete response	0	0	
Partial response	0	1	
Stable disease	2	22	
Progressive disease	24	0	
Not evaluable	3	8	

[a] Calculated by χ^2 test.

[b] According to Ref. 23.

[c] Intensity of the corresponding peak measured by quantitative mass spectrometry.

[d] Evaluated after the first two cycles of gemcitabine monotherapy.

successfully applied MALDI MS-based protein profiling techniques for predicting the efficacy of chemoradiotherapy and molecular targeting therapy (14, 15). Two-dimensional image converted analysis of liquid chromatography and mass spectrometry (2DICAL)[1] is a new LC-MS-based proteomics platform that was developed in our laboratory (16). 2DICAL can quantify protein content accurately across a theoretically unlimited number of samples without isotope labeling and thus has considerable advantages over conventional LC-MS-based methods for clinical studies (17). The predictive biomarker protein for hematologic toxicity described above was identified using 2DICAL (8).

It has been generally accepted that tumor responses do not always correlate with the outcome of patients (10, 18, 19). The rates of complete and partial responses (Response Evaluation Criteria in Solid Tumors guideline) to gemcitabine mono-

[1] The abbreviations used are: 2DICAL, two-dimensional image converted analysis of liquid chromatography and mass spectrometry; AIC, Akaike's information criterion; CC, correlation coefficient; CI, confidence interval; CV, coefficient of variance; ECOG, Eastern Cooperative Oncology Group; NCC, National Cancer Center; ID, identification; FDR, false discovery rate.

therapy are limited to ~10% (20–22), and the majority of pancreatic cancers do not show significant tumor regression. Given that the ultimate goal of gemcitabine therapy for pancreatic cancer is to achieve prolonged survival, it would be desirable to stratify patients according to survival rather than tumor response (9). In the present study, using 2DICAL, we compared the base-line plasma proteome of two extreme populations of patients who had shown distinct clinical courses after identical gemcitabine treatment.

EXPERIMENTAL PROCEDURES

Patients—Samples were collected from a total of 304 patients who had all been included in our previous study (8). All patients had metastatic (stage IVb; $n = 285$) or locally advanced (stage IVa; $n = 19$) (23) histologically or cytologically proven pancreatic ductal adenocarcinoma and had received at least two cycles of gemcitabine monotherapy (1,000 mg/m^2 intravenously over 30 min on days 1, 8, and 15 of a 28-day cycle). Two hundred eighty-one patients (92%) received gemcitabine as a first line therapy (supplemental Table S1).

Two hundred sixty-two patients (86%) were treated consecutively at the National Cancer Center (NCC) Hospital (Tokyo, Japan) between September 2002 and June 2007, and 42 patients (14%) were treated consecutively at the NCC Hospital East (Kashiwa, Japan) between September 2002 and July 2004. Survival times were determined as of May 2008. During this period, 248 patients (82%) died, and 56 patients (18%) were censored. Tumor response was evaluated after the first two cycles of gemcitabine using the Response Evaluation Criteria in Solid Tumors guideline.

Sample Preparation—Blood was collected before the first administration of gemcitabine. Plasma or serum was separated by centrifugation at 1,050 × *g* for 10 min at 4 °C and frozen until analysis as reported previously (8, 24). Macroscopically hemolyzed samples were excluded from the current analysis. Two hundred fifty-two plasma samples (83%) were collected from the NCC Hospital and Hospital East, and 52 serum samples (17%) were collected from the NCC Hospital. Written informed consent was obtained from all patients before blood sampling. The protocol of this retrospective study was reviewed and approved by the institutional ethics committee boards of the NCC (Tokyo, Japan) and the National Institute of Health Sciences (Tokyo, Japan).

LC-MS—Samples were blinded, randomized, and passed through an IgY-12 High Capacity Spin Column (Beckman Coulter, Fullerton, CA) in accordance with the manufacturer's instructions. The flow-through portion was digested with sequencing grade modified trypsin (Promega, Madison, WI) and analyzed in triplicate using a nanoflow high performance LC system (NanoFrontier nLC, Hitachi High Technologies, Tokyo, Japan) connected to an electrospray ionization quadrupole time-of-flight mass spectrometer (Q-Tof Ultima, Waters, Milford, MA). LC-MS run order was also randomized to eliminate any potential bias.

MS peaks were detected, normalized, and quantified using the in-house 2DICAL software package as described previously (16). A serial identification (ID) number was applied to each of the MS peaks detected (1–45,277). The stability of LC-MS was monitored by calculating the correlation coefficient (CC) and coefficient of variance (CV) of every triplicate measurement. The mean CC and CV ± S.D. for all 45,277 peaks observed in the 60 triplicate runs were as high as 0.970 ± 0.022 and as low as 0.056 ± 0.017, respectively.

Protein Identification by MS/MS—Peak lists were generated using the Mass Navigator software package (version 1.2) (Mitsui Knowledge Industry, Tokyo, Japan) and searched against the NCBInr database (downloaded on May 20, 2008) using the Mascot software package (version 2.2.1) (Matrix Science, London, UK). The search parameters used were as follows. A database of human proteins was selected.

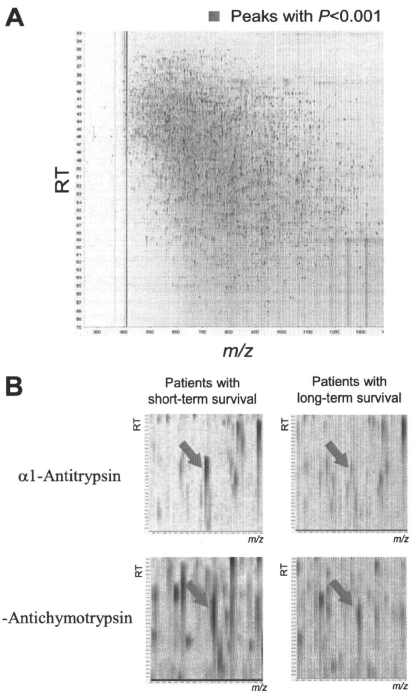

FIG. 1. *A*, two-dimensional display of all (>45,000) the MS peaks. The 637 MS peaks whose mean intensity differed significantly between patients with short term and long term survival ($p < 0.001$, Welch's *t* test) are highlighted in *red*. *B*, two MS peaks with the smallest *p* values (*upper*, $p = 2.57 \times 10^{-4}$; *bottom*, $p = 5.03 \times 10^{-4}$) in representative patients with short term (*left*) and long term (*right*) survival. *RT*, retention time.

Trypsin was designated as the enzyme, and up to one missed cleavage was allowed. Mass tolerances for precursor and fragment ions were ±2.0 and ±0.8 Da, respectively. The score threshold was set to $p < 0.05$ based on the size of the database used in the search. If a peptide matched to multiple proteins, the protein name with the highest Mascot score was selected.

Western Blot Analysis—Primary antibodies used were rabbit polyclonal antibody against human α_1-antitrypsin (Dako, Glostrup, Denmark), rabbit polyclonal antibody against human α_1-antichymotrypsin (Dako), and mouse monoclonal antibody against human complement C3b-α (Progen, Heidelberg, Germany). Ten microliters of partitioned sample was separated by SDS-PAGE and electroblotted onto a polyvinylidene difluoride membrane. The membrane was then incubated with the primary antibody and subsequently with the relevant horseradish peroxidase-conjugated anti-rabbit or anti-mouse IgG as described previously (25, 26). Blots were developed using an ECL detection system (GE Healthcare).