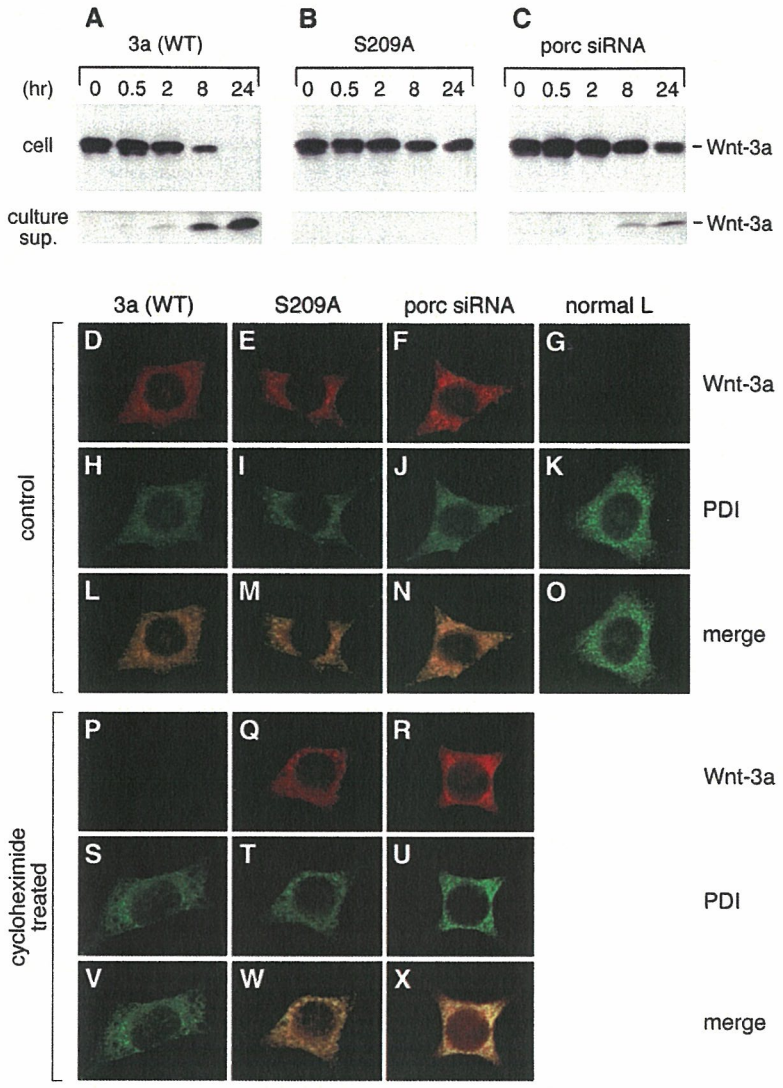


Figure 5. Defective Acylation Results in Preferential Retention of Wnt-3a Protein in the Endoplasmic Reticulum

(A–C) L cells expressing (A) wild-type Wnt-3a or (B) Wnt-3a (S209A-cl 1) as well as (C) wild-type Wnt-3a-expressing L cells defective in Porc activity (si21-6) were incubated with cycloheximide. The amounts of Wnt-3a remaining within cells ([A]–[C], upper panel) and released into the culture supernatant ([A]–[C], lower panel) were examined by immunoblotting at the indicated time points after the addition of cycloheximide.

(D–X) L cells expressing (D, H, L, P, S, and V) wild-type Wnt-3a or the (E, I, M, Q, T, and W) S209A mutant form, as well as (F, J, N, R, U, and X) wild-type Wnt-3a-expressing L cells defective in Porc activity and (G, K, and O) control L cells, were fixed 24 hr after incubation (D–O) without or (P–X) with cycloheximide. Cells were then stained with anti-Wnt-3a antibody (shown as red: [D]–[G] and [P]–[R]) or anti-PDI antibody, which is an ER marker (shown as green: [H]–[K] and [S]–[U]). (L–O and V–X) Merged pictures are also shown. The percentage of stained cells is shown in Table S2.

porc-deficient cells treated with cycloheximide (Figures 4C and 5C), although reduced levels of Wnt-3a secretion were detected, reflecting residual Porc activity. These results show that *porc* is required for secretion, as well as for Ser209-dependent acylation, of Wnt-3a from the cultured cells.

Wnt-3a with Defective Acylation Is Not Transferred from the ER

To determine the importance of acylation to Wnt-3a secretion, we examined the intracellular localization of Wnt protein with defective acylation (Figures 5D–5X; Table S2). Most Wnt-3a protein transcribed from the transfected gene is retained in the ER, probably due to limitations of protein trafficking from the ER (Figures 5D–5O). During cycloheximide-mediated inhibition of de novo protein synthesis, S209A protein was retained in the ER (Figures 5Q and 5W), whereas most newly synthesized wild-type Wnt-3a protein was secreted from cells under the same conditions (Figures 5P and 5V). Similarly, under the same conditions, most Wnt-3a protein was retained in the ER in *porc*-deficient cells (Figures 5R and 5X). Thus, Ser209-dependent acylation appears to be required for transit of Wnt-3a from the ER.

Ser209 Is Required for the Function and Appropriate Localization of Wnt-3a in *Xenopus* Embryos

To examine whether Ser209-dependent acylation would also be required for the function and secretion of Wnt-3a in vivo, we next injected S209A mRNA into *Xenopus* embryos (Figure 6). While the injection of wild-type Wnt-3a or S211A mRNAs into the ventral side of *Xenopus* eggs at the four (or two)-cell stage caused ectopic axis formation (Figures 6B and 6D; Table S3), as well as induction of the target genes *Siamois* and *Xnodal related-3* (Figure 6G), injection of S209A mRNA had no obvious effect, indicating that Ser209 is essential for the in vivo function of Wnt-3a (Figures 6C and 6G). Furthermore, whereas most wild-type Wnt-3a and S211A proteins were transported to the apical cell border, most S209A proteins were retained inside the expressing cells, probably in the ER (Figures 6I–6K). These results confirm that Ser209, which is required for acylation and secretion of Wnt-3a in cultured cells, is also essential for proper intracellular transport of Wnt-3a proteins in vivo.

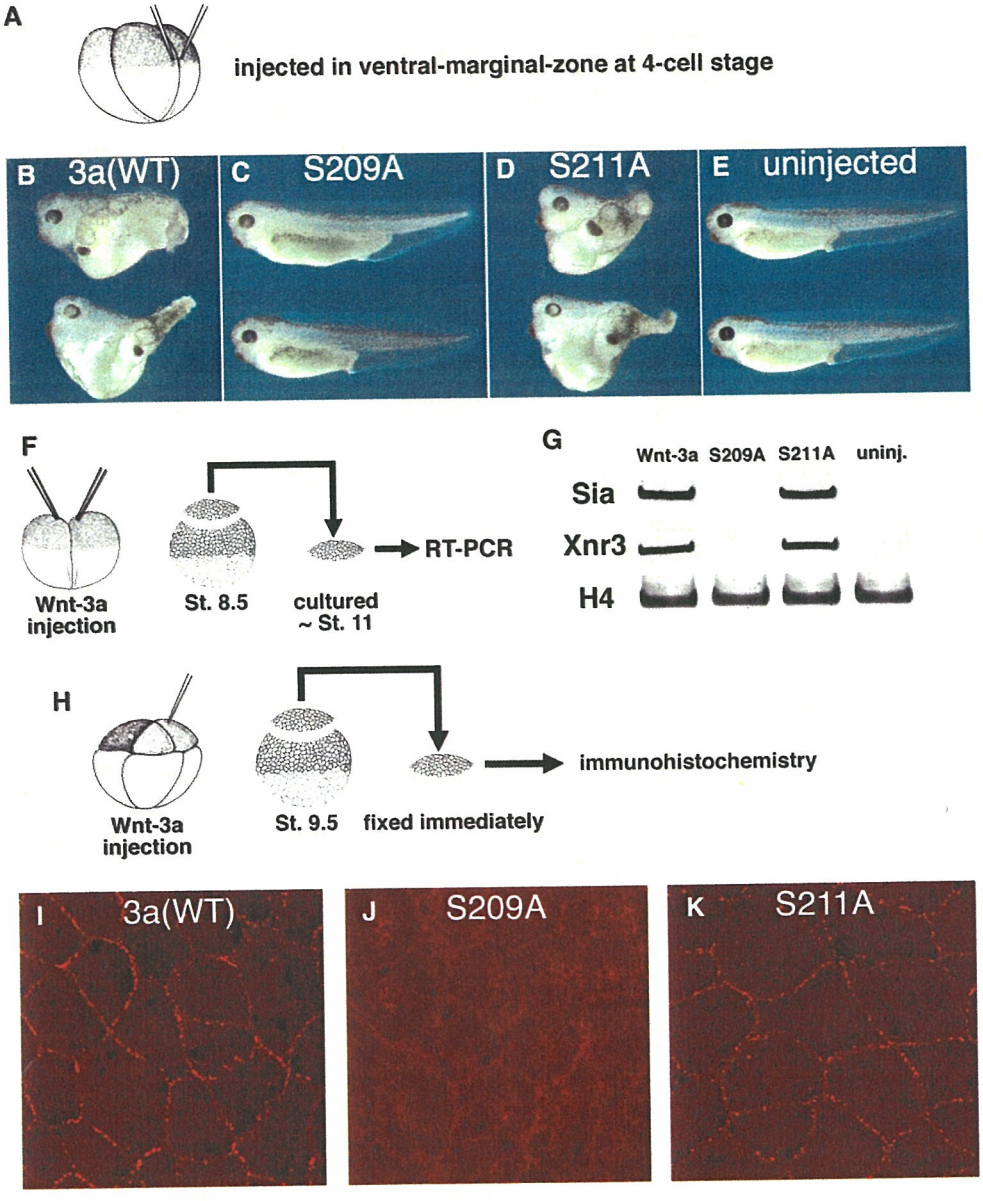


Figure 6. Acylation-Defective Wnt-3a Protein Is Largely Retained within the Cells of *Xenopus* Embryos

(A–K) (A), (F), and (H) indicate schematic representations of the experiments performed in (B)–(E), (G), and (I)–(K), respectively. (B–E) *Xenopus* embryos injected with mRNAs encoding (B) wild-type, (C) S209A, or (D) S211A forms of Wnt-3a, as well as (E) uninjected embryos, are shown. (G) Expression of *siamois* (Sia), *Xnodal related-3* (Xnr3), and *histone H4* (H4) in animal caps dissected from embryos injected with mRNAs encoding these forms of Wnt-3a, as well as in those from uninjected embryos, was analyzed by use of the reverse transcription-polymerase chain reaction. (I–K) Confocal images of animal caps dissected from the embryos injected with mRNAs encoding (I) wild-type, (J) S209A, or (K) S211A forms of Wnt-3a stained with anti-Wnt-3a antibody are shown.

Discussion

In this study, we demonstrated that Wnt-3a protein is acylated with an unsaturated fatty acid, palmitoleic acid (C16:1), at a conserved serine at the 209th residue, in addition to the previously reported Cys77 palmitoylation. We analyzed a number of mutant forms of Wnt-3a to demonstrate that Ser209 is required for appropriate trafficking of Wnt-3a protein from the ER during secretion, both in cultured cells and in embryos. In addition, Ser209-dependent acylation requires the function of Porc, a putative membrane-bound O-acyltransferase, in the ER. Thus, these results strongly suggest that Ser209-dependent acylation, catalyzed by Porc in the ER, is essential for transport of Wnt-3a from the ER during the secretion process.

Since acylation has been implicated in regulation of protein trafficking to intracellular organelles and particular domains of the plasma membrane (Smotrys and Linder, 2004; Huang and El-Husseini, 2005; Resh, 1999), Ser209-dependent acylation may be required for targeting of Wnt-3a proteins to specific organelles or membrane components required for secretion. Consistent with this model, an inhibitor of acyltransferase activity is known to inhibit the intracellular trafficking of Wg protein (Zhai et al., 2004). On the other hand, since

Figure 7. The Amino Acid Sequence Surrounding Ser209 Is Highly Conserved among Members of the Wnt Family

Aligned amino acid sequences surrounding Ser209 are shown for various forms of Wnt protein. Ser209 is shown in red, and other conserved residues are shown in blue. The consensus sequence among the members of the Wnt family is indicated at the bottom.

Wnt protein is transported between cells by lipoprotein particles in *Drosophila* wing imaginal discs (Greco et al., 2001; Panakova et al., 2005), an acylation-dependent trafficking system might be prerequisite for attachment of Wnt proteins to lipoprotein particles for secretion. Although it is not certain at present whether Wnt protein is secreted from cultured mammalian cells in association with lipoprotein particles, further characterization of the secreted Wnt protein might reveal whether acylation of Wnt protein has this proposed role in mammalian cell culture systems.

In addition, acylation may affect the function of several molecules involved in Wnt secretion, such as Wntless/Evi, a seven-pass membrane protein required for Wnt secretion in *Drosophila* and mammalian cultured cells (Banziger et al., 2006; Bartscherer et al., 2006), as well as the retromer complex, a protein complex involved in intracellular membrane trafficking and required for the long-range extracellular transport of Wnt protein in *C. elegans* (Coudreuse et al., 2006; Prasad and Clark, 2006). Because Wntless/Evi is primarily localized at the Golgi apparatus or plasma membrane, and since components of the retromer complex can be identified in *trans*-Golgi and trafficking vesicles, acylation-dependent trafficking of Wnt-3a from the ER appears to be prerequisite for Wntless/Evi and retromer complex function.

Another important aspect of this study pertains to the discovery of an anticipated lipid modification. This modification demonstrates several important characteristics, especially as related to the type of fatty acids involved. We demonstrated modification of Wnt-3a protein at Ser209 with monounsaturated fatty acids, such as palmitoleic acid (C16:1), but not with saturated fatty acids like palmitic acid (C16:0). Since the addition of deuterium-labeled palmitic acid (C16:0) into the cell culture medium resulted in a partial shift of the isotopic ion distribution of palmitoleic acid (C16:1) in mass spectrometry, Wnt-3a protein appeared to be modified with C16:1-fatty acid metabolically processed from palmitic acid (C16:0). Most cases of acylation with unsaturated fatty acids occur at sites where acylation with saturated fatty acids also occurs (DeMar et al., 1999; Dizhoor et al., 1992; Johnson et al., 1994; Kokame et al., 1992; Liang et al., 2001; Muszbek and Laposata, 1993; Neubert et al., 1992). In contrast, the results of our mass spectrometry analysis indicate a strong preference for modification with monounsaturated fatty acids at Ser209 of Wnt-3a. Given that cells store considerable amounts of saturated fatty acids as substrates for acyltransferases, the predominant modification with a palmitoleoyl (C16:1) moiety must be the result of preferential utilization of this unsaturated fatty acid as a substrate.

In addition to modification with a palmitoleoyl (C16:1) moiety, another interesting feature of Ser209 acylation involves the amino acid residue to which the acyl moiety is attached. In most cases of protein fatty acylation examined to date, including N-myristoylation, S-palmitoylation, and N-palmitoylation, a serine residue is not utilized for connecting to the acyl moieties, although there are a few exceptions, including attachment of oxyester-linked octanoate (C8:0) to serine in ghrelin, a growth-hormone-releasing peptide of 28 amino acids (Kojima et al., 1999; Smotrys and Linder, 2004). Here, we showed a different example of serine-linked O-acylation, where serine is used to bond a monounsaturated acyl moiety to a protein. Thus, Wnt-3a acylation demonstrates several unusual characteristics, suggesting that some specific machinery appears to be involved in this acylation.

One strong candidate for a player involved in this machinery is Porc. Since most Wnt acylation detected with the labeling assay of this study was dependent on Ser209 and was abolished in cells with markedly reduced Porc activity, *porc* appears to be required for acylation at Ser209. If this is the case, it will be important to elucidate the mechanism by which this enzyme utilizes unsaturated fatty acids and recognizes serine residues. Interestingly, additional members of the MBOAT family of membrane-bound O-acyltransferases, including acyl-CoA cholesterol acyltransferase (ACAT) and diacylglycerol acyltransferase (DGAT), utilize monounsaturated fatty acids as substrates (Cases et al., 1998; Seo et al., 2001). Thus, certain structural characteristics of this family may contribute to their preferential utilization of unsaturated fatty acids, as well as their role in O-acylation.

Another question pertains to the biological significance of acylation by unsaturated fatty acids. Although the functional significance of such acylation remains unclear, the bent structure produced by the double bond as a result of this process may influence the interaction of acylated proteins with lipid structures. Both in vivo and in vitro results indicate that acylation with unsaturated fatty acids results in displacement of proteins, including Fyn, annexinII, and Gai, from membrane domains within ordered lipid structures (Liang et al., 2001; Moffett et al., 2000; Zhao and Hardy, 2004). Thus, it seems unlikely that unsaturated fatty acylation enhances protein targeting into some ordered lipid structures, but rather that the folded structure might be advantageous for packaging fatty acid chains into the interior of small lipid particles. As such, we can speculate that palmitoleoyl (C16:1) modification may enable packaging of Wnt proteins into protein-lipid particles. An intracellular precursor of a lipoprotein particle might be a candidate for such a particle (Greco et al., 2001; Panakova et al., 2005).

Finally, it is of interest to note that the amino acid sequence surrounding Ser209 (C-K-C-H-G-(LIVMT)-S-G-S-C, where the bold S indicates Ser209) is highly conserved among members of the Wnt family, although the function of this motif remains unknown (Figure 7). Our results indicate that Ser209 in this conserved motif is essential for palmitoleoyl (C16:1) modification and secretion of Wnt-3a protein, suggesting that other Wnt members may likewise be modified by palmitoleoyl acid (C16:1), depending on the presence of this conserved motif. An attempt to reveal the role of this motif might provide clues regarding the molecular mechanism behind acylation.

Experimental Procedures

Cell Culture, Transfection, and Metabolic Labeling of Wnt Protein

L cells were cultured at 37°C in a 1:1 mixture of DMEM and Ham's F12 medium supplemented with 8.3% fetal calf serum and antibiotics. Stable transfectants expressing truncated or point-mutated forms of Wnt-3a were established as previously described (Shibamoto et al., 1998). Supernatants from cultures of Wnt-3a- or Wnt-5a-producing L cells were prepared as previously described (Shibamoto et al., 1998; Yamanaka et al., 2002). For inhibition of de novo protein synthesis, L cells were treated with 10 μg/ml cycloheximide for up to 24 hr; for that of Porc activity, a plasmid vector, pSilencer 3.0-H1 (Ambion), expressing siRNA specific for *porc* (si21: 5′-AAGTT GTCACAAGCTGGAACC-3′), was used to transfect Wnt-3a-expressing L cells. Stable transfectants with varying levels of defective *porc* expression were thus established and used in the experiments.

L cells secreting various forms of Wnt-3a (or Wnt-5a) and control L cells were inoculated into 35 mm dishes and were then incubated overnight at 37°C. Next, these cells were incubated for 36 hr in serum-free medium containing 0.37 MBq/ml [U-^{14}C] palmitic acid (30.4 GBq/mmol) or 14.8 MBq/ml [9,10-^{3}H] palmitic acid (2.2 TBq/mmol), or they were incubated in serum-free medium without labeled fatty acid for 20 hr and then with 0.37 MBq/ml [U-^{14}C] palmitic acid (30.4 GBq/mmol) for 4 hr. The culture supernatant or cell lysate was collected and immunoprecipitated with anti-Wnt-3a (or anti-Wnt-5a) antibody (R.T. and S.T., unpublished data) and protein G Sepharose beads. Immunoprecipitates mixed with sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer containing 2-mercaptoethanol at the final concentration of 1.3 M were incubated at 37°C for 1 hr, separated by SDS-polyacrylamide gel electrophoresis, and visualized by autoradiography with a Kodak BioMax TranScreen LE system or by image analysis with a BAS 2500 (Fuji Photo Film).

Antibodies, Immunoblotting, and Immunocytology

The preparation of monoclonal anti-Wnt-3a and anti-Wnt-5a antibodies is described elsewhere (R.T and S.T., unpublished data). These antibodies specifically recognize Wnt-3a or Wnt-5a proteins, respectively. Anti-protein disulfide isomerase rabbit polyclonal antibody (Calbiochem), anti-HA antibody (Covance), and anti-FLAG M2 affinity gel (Sigma), were purchased. Immunoblotting was performed according to a standard protocol. For immunocytological analysis, cells were fixed with 3% paraformaldehyde in PBS(−) for 10 min at room temperature, made permeable by treatment with cold 100% methanol for 10 min, and then incubated with anti-Wnt-3a antibody and anti-protein disulfide isomerase (anti-PDI, a rabbit polyclonal antibody) for 1 hr at room temperature. Thereafter, the cells were probed with CY3-conjugated anti-mouse IgG and Alexa 488-conjugated anti-rabbit secondary antibodies and were observed under a confocal microscope (Carl-Zeiss, LSM510).

Nano-Flow Liquid Chromatography

FLAG-tagged Wnt-3a protein was purified with anti-FLAG M2 antibody from the culture supernatant of Wnt-3a-expressing L cells in serum-free medium. Purified Wnt-3a protein was reduced with 0.2 M dithiothreitol, and cysteine alkylation was carried out with monoacrylamide. Wnt-3a protein was directly applied to a gel for 8.0% SDS-PAGE, after which the gel was stained with silver-staining reagents or Coomassie brilliant blue (CBB). The protein band at ~45 kDa was excised from a strip of the gel. The excised protein band that had been stained with silver-staining reagents was destained with 20 mM EDTA 2Na-50 mM NH_4HCO_3 in 30% aqueous acetonitrile and was then washed with 15 mM potassium hexacyanoferrate(III)-50 mM sodium thiosulfate. The band stained with CBB was destained with 100 mM NH_4HCO_3 in 60% aqueous acetonitrile. The gel was then subjected to in-gel digestion with trypsin (Promega).

The trypsin digest was extracted from the gel with 0.1% trifluoroacetic acid in 60% aqueous acetonitrile. The resultant solution was injected into an Ultimate nano-LC system (Dionex), where the digested peptides were first concentrated with a C18 trapping column (0.3 mm × 1 mm, Dionex, Idstein, Germany) at a flow rate of 30 μl/min, and then separated by using a C_{18}-Pepmap column (0.075 × 150 mm, Dionex). A linear gradient of solvent A (0.1% trifluoroacetic acid in water) and solvent B (0.1% trifluoroacetic acid in acetonitrile) was used for the separation, and the peptides were eluted by increasing the concentration of solvent B from 5% to 80% over a period of 60 min at a flow rate of 200 nl/min. The effluent was monitored at 214 and 280 nm and was directly blotted at 1 min intervals onto the flat surface of a stainless steel plate (a MALDI sample plate) over a 96 min period. Thereafter, the matrix solution (5 mg/ml α-cyano-4-hydroxycinnamic acid) was blotted manually onto each sample spot and then dried.

Matrix-Assisted Laser Desorption/Ionization, MALDI, Mass Spectrometry

Overall peptide identification was carried out by using a MALDI-TOF/TOF (4700 proteomics analyzer, Applied Biosystems, Framingham, MA), followed by a database search with Mascot ver. 2.0 (Matrix Science, Manchester, UK). Ions were generated by irradiating the sample area with a 200 Hz Nd:YAG laser operated at 355 nm. Calibration was performed by using MH^+ ions from a mixture of angiotensin I (m/z 1296.6), dynorphin (m/z 1604.0), ACTH (1-24) (m/z 2932.6), and β-endorphin (m/z 3463.8). The isotopic envelopes of the observed ions were closely compared with the theoretical envelopes generated by *Isotopica*, a software aid for calculating and assessing complex isotopic envelopes (http://coco.protein.osaka-u.ac.jp/Isotopica) (Fernandez-de-Cossio et al., 2004). For MS/MS, the precursor ions were accelerated at 8 kV in MS1 and were fragmented in a collision cell by using air as the collision gas. The resultant fragment ions, re-accelerated at 15 kV, were analyzed in MS2 equipped with a reflectron. MS/MS spectra were interpreted by *SeqMS*, a software aid for de novo sequencing by MS/MS (http://www.protein.osaka-u.ac.jp/rcsfp/profiling) (Fernandez-de-Cossio et al., 2000).

Xenopus Injection, Detection of Target Gene Expression, and Immunohistology

For morphological analysis, in vitro-synthesized wild-type, S209A, or S211A Wnt-3a mRNA sequences were injected into the ventral marginal zones of *Xenopus* eggs at the four-cell stage (Sokol et al., 1991). To examine the expression of *Siamois* and *Xnr3* (Brannon and Kimelman, 1996; McKendry et al., 1997), we injected the various Wnt-3a mRNA sequences at the two-cell stage, after which the animal caps of the injected embryos were dissected at stage 8.5, cultured to stage 11, and analyzed for gene expression by use of the reverse transcription-polymerase chain reaction. The following specific primers were used: forward, 5′-GATAACTGGCATTCCTGAGC-3′; reverse, 5′-ACAAGTCAGTGTGGTGATTC-3′ (23 cycles) for *siamois*; forward, 5′-CCATGTGAGCACCGTTCC-3′; reverse, 5′-GAGCAAACT CTTAATGTAG-3′ (18 cycles) for *Xnr3*; and forward, 5′-ATAACATC CAGGGCATCACC-3′; reverse, 5′-ACATCCATAGCGGTGACGGT-3′ (18 cycles) for *histone H4* as the internal input control. For immunohistological analysis, the various Wnt-3a mRNA sequences were injected into the blastomeres in the animal side of eight-cell-stage blastulas, and the explants from the injected embryos were fixed at stage 9.5 in MEMFA (0.1 M MOPS [pH 7.4], containing 2 mM EGTA, 1 mM $MgSO_4$, and 3.7% formaldehyde) for 1 hr at room temperature. After bleaching in 10% H_2O_2 and blocking with FCS-PBS (PBS with 15% fetal calf serum) for 20 min, the explants were incubated overnight at 4°C with anti-Wnt-3a antibody in FCS-PBS. They were then incubated for 2 hr at room temperature with rhodamine-conjugated

rabbit anti-mouse IgG. The washed explants were mounted and observed under a confocal microscope (Carl-Zeiss, LSM510).

Supplemental Data

Supplemental Data include three figures and three tables and are available at http://www.developmentalcell.com/cgi/content/full/11/6/791/DC1/.

Acknowledgments

We would like to thank Dr. N. Kinoshita, Dr. M. Ohashi, Ms. K. Irie, and all members of the Takada and Ueno labs for their technical support and helpful discussion. We also thank Ms. A. Taya (Institute for Protein Research, Osaka University) for supporting our liquid chromatography and mass spectrometry analyses. This work was supported by a grant-in-aid for scientific research from the Japanese Ministry of Education, Science, Culture, and Sports (17082012 to S.T. and 15GS0320 to T.T.) and by grants from the Japan Science and Technology Corporation and Mitsubishi Foundation to S.T.

Received: May 9, 2006
Revised: August 31, 2006
Accepted: October 4, 2006
Published: December 4, 2006

References

Banziger, C., Soldini, D., Schutt, C., Zipperlen, P., Hausmann, G., and Basler, K. (2006). Wntless, a conserved membrane protein dedicated to the secretion of Wnt proteins from signaling cells. Cell *125*, 509–522.

Bartscherer, K., Pelte, N., Ingelfinger, D., and Boutros, M. (2006). Secretion of Wnt ligands requires Evi, a conserved transmembrane protein. Cell *125*, 523–533.

Bizzozero, O.A. (1995). Chemical analysis of acylation sites and species. Methods Enzymol. *250*, 361–379.

Brannon, M., and Kimelman, D. (1996). Activation of Siamois by the Wnt pathway. Dev. Biol. *180*, 344–347.

Cases, S., Smith, S.J., Zheng, Y.W., Myers, H.M., Lear, S.R., Sande, E., Novak, S., Collins, C., Welch, C.B., Lusis, A.J., et al. (1998). Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc. Natl. Acad. Sci. USA *95*, 13018–13023.

Coudreuse, D.Y., Roel, G., Betist, M.C., Destree, O., and Korswagen, H.C. (2006). Wnt gradient formation requires retromer function in Wnt-producing cells. Science *312*, 921–924.

DeMar, J.C., Jr., Rundle, D.R., Wensel, T.G., and Anderson, R.E. (1999). Heterogeneous N-terminal acylation of retinal proteins. Prog. Lipid Res. *38*, 49–90.

Dizhoor, A.M., Ericsson, L.H., Johnson, R.S., Kumar, S., Olshevskaya, E., Zozulya, S., Neubert, T.A., Stryer, L., Hurley, J.B., and Walsh, K.A. (1992). The NH2 terminus of retinal recoverin is acylated by a small family of fatty acids. J. Biol. Chem. *267*, 16033–16036.

Fernandez-de-Cossio, J., Gonzalez, J., Satomi, Y., Shima, T., Okumura, N., Besada, V., Betancourt, L., Padron, G., Shimonishi, Y., and Takao, T. (2000). Automated interpretation of low-energy collision-induced dissociation spectra by SeqMS, a software aid for de novo sequencing by tandem mass spectrometry. Electrophoresis *21*, 1694–1699.

Fernandez-de-Cossio, J., Gonzalez, L.J., Satomi, Y., Betancourt, L., Ramos, Y., Huerta, V., Amaro, A., Besada, V., Padron, G., Minamino, N., and Takao, T. (2004). Isotopica: a tool for the calculation and viewing of complex isotopic envelopes. Nucleic Acids Res. *32*, W674–W678.

Greco, V., Hannus, M., and Eaton, S. (2001). Argosomes: a potential vehicle for the spread of morphogens through epithelia. Cell *106*, 633–645.

Hacker, U., Nybakken, K., and Perrimon, N. (2005). Heparan sulphate proteoglycans: the sweet side of development. Nat. Rev. Mol. Cell Biol. *6*, 530–541.

Hofmann, K. (2000). A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling. Trends Biochem. Sci. *25*, 111–112.

Huang, K., and El-Husseini, A. (2005). Modulation of neuronal protein trafficking and function by palmitoylation. Curr. Opin. Neurobiol. *15*, 527–535.

Johnson, D.R., Bhatnagar, R.S., Knoll, L.J., and Gordon, J.I. (1994). Genetic and biochemical studies of protein N-myristoylation. Annu. Rev. Biochem. *63*, 869–914.

Kadowaki, T., Wilder, E., Klingensmith, J., Zachary, K., and Perrimon, N. (1996). The segment polarity gene porcupine encodes a putative multitransmembrane protein involved in Wingless processing. Genes Dev. *10*, 3116–3128.

Kojima, M., Hosoda, H., Date, Y., Nakazato, M., Matsuo, H., and Kangawa, K. (1999). Ghrelin is a growth-hormone-releasing acylated peptide from stomach. Nature *402*, 656–660.

Kokame, K., Fukada, Y., Yoshizawa, T., Takao, T., and Shimonishi, Y. (1992). Lipid modification at the N terminus of photoreceptor G-protein α-subunit. Nature *359*, 749–752.

Lewis, P.M., Dunn, M.P., McMahon, J.A., Logan, M., Martin, J.F., St-Jacques, B., and McMahon, A.P. (2001). Cholesterol modification of sonic hedgehog is required for long-range signaling activity and effective modulation of signaling by Ptc1. Cell *105*, 599–612.

Liang, X., Nazarian, A., Erdjument-Bromage, H., Bornmann, W., Tempst, P., and Resh, M.D. (2001). Heterogeneous fatty acylation of Src family kinases with polyunsaturated fatty acids regulates raft localization and signal transduction. J. Biol. Chem. *276*, 30987–30994.

Lin, X. (2004). Functions of heparan sulfate proteoglycans in cell signaling during development. Development *131*, 6009–6021.

Linder, M.E., and Deschenes, R.J. (2003). New insights into the mechanisms of protein palmitoylation. Biochemistry *42*, 4311–4320.

Logan, C.Y., and Nusse, R. (2004). The Wnt signaling pathway in development and disease. Annu. Rev. Cell Dev. Biol. *20*, 781–810.

Mann, R.K., and Beachy, P.A. (2004). Novel lipid modifications of secreted protein signals. Annu. Rev. Biochem. *73*, 891–923.

McKendry, R., Hsu, S.C., Harland, R.M., and Grosschedl, R. (1997). LEF-1/TCF proteins mediate wnt-inducible transcription from the *Xenopus* nodal-related 3 promoter. Dev. Biol. *192*, 420–431.

Miura, G.I., Buglino, J., Alvarado, D., Lemmon, M.A., Resh, M.D., and Treisman, J.E. (2006). Palmitoylation of the EGFR ligand Spitz by Rasp increases Spitz activity by restricting its diffusion. Dev. Cell *10*, 167–176.

Moffett, S., Brown, D.A., and Linder, M.E. (2000). Lipid-dependent targeting of G proteins into rafts. J. Biol. Chem. *275*, 2191–2198.

Moon, R.T., Kohn, A.D., De Ferrari, G.V., and Kaykas, A. (2004). WNT and β-catenin signalling: diseases and therapies. Nat. Rev. Genet. *5*, 691–701.

Muszbek, L., and Laposata, M. (1993). Myristoylation of proteins in platelets occurs predominantly through thioester linkages. J. Biol. Chem. *268*, 8251–8255.

Neubert, T.A., Johnson, R.S., Hurley, J.B., and Walsh, K.A. (1992). The rod transducin α subunit amino terminus is heterogeneously fatty acylated. J. Biol. Chem. *267*, 18274–18277.

Nusse, R. (2003). Wnts and Hedgehogs: lipid-modified proteins and similarities in signaling mechanisms at the cell surface. Development *130*, 5297–5305.

Panakova, D., Sprong, H., Marois, E., Thiele, C., and Eaton, S. (2005). Lipoprotein particles are required for Hedgehog and Wingless signalling. Nature *435*, 58–65.

Pepinsky, R.B., Zeng, C., Wen, D., Rayhorn, P., Baker, D.P., Williams, K.P., Bixler, S.A., Ambrose, C.M., Garber, E.A., Miatkowski, K., et al. (1998). Identification of a palmitic acid-modified form of human Sonic hedgehog. J. Biol. Chem. *273*, 14037–14045.

Porter, J.A., Ekker, S.C., Park, W.J., von Kessler, D.P., Young, K.E., Chen, C.H., Ma, Y., Woods, A.S., Cotter, R.J., Koonin, E.V., and Beachy, P.A. (1996a). Hedgehog patterning activity: role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain. Cell *86*, 21–34.

Porter, J.A., Young, K.E., and Beachy, P.A. (1996b). Cholesterol modification of hedgehog signaling proteins in animal development. Science *274*, 255–259.

Prasad, B.C., and Clark, S.G. (2006). Wnt signaling establishes anteroposterior neuronal polarity and requires retromer in *C. elegans*. Development *133*, 1757–1766.

Resh, M.D. (1999). Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. Biochim. Biophys. Acta. *1451*, 1–16.

Reya, T., and Clevers, H. (2005). Wnt signalling in stem cells and cancer. Nature *434*, 843–850.

Seo, T., Oelkers, P.M., Giattina, M.R., Worgall, T.S., Sturley, S.L., and Deckelbaum, R.J. (2001). Differential modulation of ACAT1 and ACAT2 transcription and activity by long chain free fatty acids in cultured cells. Biochemistry *40*, 4756–4762.

Shibamoto, S., Higano, K., Takada, R., Ito, F., Takeichi, M., and Takada, S. (1998). Cytoskeletal reorganization by soluble Wnt-3a protein signalling. Genes Cells *3*, 659–670.

Smotrys, J.E., and Linder, M.E. (2004). Palmitoylation of intracellular signaling proteins: regulation and function. Annu. Rev. Biochem. *73*, 559–587.

Sokol, S., Christian, J.L., Moon, R.T., and Melton, D.A. (1991). Injected Wnt RNA induces a complete body axis in *Xenopus* embryos. Cell *67*, 741–752.

Tanaka, K., Kitagawa, Y., and Kadowaki, T. (2002). *Drosophila* segment polarity gene product porcupine stimulates the posttranslational N-glycosylation of wingless in the endoplasmic reticulum. J. Biol. Chem. *277*, 12816–12823.

Thorpe, C.J., Schlesinger, A., Carter, J.C., and Bowerman, B. (1997). Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. Cell *90*, 695–705.

van den Heuvel, M., Harryman-Samos, C., Klingensmith, J., Perrimon, N., and Nusse, R. (1993). Mutations in the segment polarity genes wingless and porcupine impair secretion of the wingless protein. EMBO J. *12*, 5293–5302.

Willert, K., Brown, J.D., Danenberg, E., Duncan, A.W., Weissman, I.L., Reya, T., Yates, J.R., 3rd, and Nusse, R. (2003). Wnt proteins are lipid-modified and can act as stem cell growth factors. Nature *423*, 448–452.

Yamanaka, H., Moriguchi, T., Masuyama, N., Kusakabe, M., Hanafusa, H., Takada, R., Takada, S., and Nishida, E. (2002). JNK functions in the non-canonical Wnt pathway to regulate convergent extension movements in vertebrates. EMBO Rep. *3*, 69–75.

Zhai, L., Chaturvedi, D., and Cumberledge, S. (2004). *Drosophila* wnt-1 undergoes a hydrophobic modification and is targeted to lipid rafts, a process that requires porcupine. J. Biol. Chem. *279*, 33220–33227.

Zhao, H., and Hardy, R.W. (2004). Long-chain saturated fatty acids induce annexin II translocation to detergent-resistant membranes. Biochem. J. *381*, 463–469.

Oncogene (2006) 25, 7373–7380

www.nature.com/onc

ORIGINAL ARTICLE

Suppression of tumorigenicity, but not anchorage independence, of human cancer cells by new candidate tumor suppressor gene CapG

A Watari[1], K Takaki[1], S Higashiyama[1], Y Li[1], Y Satomi[2], T Takao[2], A Tanemura[3], Y Yamaguchi[3], I Katayama[3], M Shimakage[4], I Miyashiro[5], K Takami[5], K Kodama[5] and M Yutsudo[1]

[1]*Research Institute for Microbial Diseases, Osaka University, Suita, Osaka, Japan;* [2]*Institute for Protein Research, Osaka University, Osaka, Japan;* [3]*Graduate School of Medicine, Osaka University, Osaka, Japan;* [4]*Osaka National Hospital, Chuo, Osaka, Japan and* [5]*Osaka Medical Center for Cancer and Cardiovascular Diseases, Higashinari, Osaka, Japan*

Previously, we isolated a series of cell lines from a human diploid fibroblast lineage as a model for multistep tumorigenesis in humans. After passaging a single LT-transfected fibroblast clone, differently progressed cell lines were obtained, including immortalized, anchorage-independent and tumorigenic cell lines. In the present paper, we analysed the gene expression profiles of these model cell lines, and observed that expression of the CapG protein was lost in the tumorigenic cell line. To examine the possibility that loss of CapG protein expression was required for tumorigenic progression, we transfected CapG cDNA into the tumorigenic cell line and tested for tumor-forming ability in nude mice. Results showed that ectopic expression of CapG suppressed tumorigenicity, but not growth in soft agar or liquid medium. We also found that certain cancer cell lines including stomach cancer, lung cancer and melanoma had also lost CapG expression. One such cancer cell line AZ521 also became non-tumorigenic after the introduction of CapG cDNA. Moreover, we showed that CapG expression was repressed in small-cell lung cancer tissues. Together, our findings indicated that CapG is a new tumor suppressor gene involved in the tumorigenic progression of certain cancers.
Oncogene (2006) **25,** 7373–7380. doi:10.1038/sj.onc.1209732; published online 12 June 2006

Keywords: CapG; tumor suppressor gene; multistep tumorigenesis; gastric cancer; lung cancer

Introduction

Carcinogenesis is thought to proceed in a stepwise fashion with the accumulation of multiple genetic abnormalities, such as activation of proto-oncogenes and inactivation of tumor suppressor genes (Hahn and Weinberg, 2002). In the case of colorectal cancer, for example, sequential alterations of a specific set of genes, *APC*, *K-ras* and *p53*, can account for each clinical stage of carcinogenesis (Rajogopalan *et al.*, 2003). However, the direct correlation between such genetic abnormalities and defined clinical stages and the molecular mechanism of progression through the different clinical stages remain poorly understood.

To clarify the mechanisms of multistep carcinogenesis, we decided that it would be useful to study a series of differently transformed cell lines derived from a single line of normal cells and then analyse the differences in gene expression between these cell lines. To this end, we previously isolated a series of variously transformed cell lines (retinoblastoma (RB) cell lineage) from human skin fibroblasts (RB) from a patient with hereditary RB (Oka *et al.*, 1999). Whereas RB cells had the normal diploid set of 46 chromosomes, one copy of chromosome 13 contained a large deletion spanning the region from q14 to q22. We introduced early genes of SV40 into RB cells and obtained several mortal clones with extended lifespan (RBSV). After repeated passages of a single RBSV cell clone, we succeeded in isolating immortalized (RBI), anchorage-independent (RBS) and tumorigenic (RBT) cell lines. Such model cell lines of multistep tumorigenesis are very rare, and to our knowledge no other similar cell lines have been previously published.

In the present study, we searched for differences in gene expression between these model cell lines and found that CapG protein expression was completely absent only in the most progressed line, the tumorigenic RBT cells. When introduced into RBT cells, the *CapG* gene suppressed RBT tumorigenicity in nude mice, but did not affect RBT colony formation in soft agar. This effect of CapG was not limited to our model cell lines. Although the *CapG* gene is ubiquitously expressed in normal tissues, expression was frequently lost in human cancer cell lines and tissues including the tumorigenic gastric cancer cell line AZ521. This cell line was also converted to a non-tumorigenic state by ectopic expression of CapG protein. These results suggested that the *CapG* functioned as a tumor suppressor gene and was

Correspondence: Dr M Yutsudo, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita 565-0871, Osaka, Japan.
E-mail: yutsudo@biken.osaka-u.ac.jp
Received 1 November 2005; revised 18 April 2006; accepted 5 May 2006; published online 12 June 2006

involved in tumorigenic conversion in various human cancers.

Results

CapG expression was lost at the tumorigenic stage of the RB cell lineage

As previously reported (Oka *et al.*, 1999), we isolated a set of human cell lines (RB cell lineage) as a model of multistep carcinogenesis. These cell lines were derived from a single human diploid fibroblast strain (RB) that was transformed by the introduction of SV40 early genes (RBSV) and resultant lines selected for progression to the immortalized (RBI), anchorage-independent (RBS) or tumorigenic (RBT) stages (Figure 1a). To investigate for altered gene expression between the RB cell lineage stages, we performed Western blot analysis using various antibodies. When one polyclonal antibody (HD-2) was used, a protein band with apparent molecular weight of 45 kDa was found to be absent only from the RBT tumorigenic cell line (Supplementary Figure 1a). This antibody was made by us to analyse another cancer-related protein, but the antibody obtained recognized several unknown protein bands probably because of contamination of other proteins in the antigen preparation for immunization. During analyses using this antibody, we found loss of the 45 kDa protein in some cell lines including the RBT cell line and strong expression of this protein in other cell lines (Supplementary Figure 1b). We next determined a protein spot corresponding to the 45 kDa protein on two-dimensional gel (Supplementary Figure 2a). Mass spectrometric analysis identified this protein as CapG (Supplementary Figure 2b). We then cloned CapG cDNA, expressed and purified glutathione-S-transferase-fused CapG protein from *Escherichia coli* and generated polyclonal anti-CapG antibody (Supplementary Figure 3). Using this antibody, we confirmed that RBT cells showed no CapG protein expression (Figure 1b). To understand the mechanism by which CapG protein expression was inhibited in the RBT cell line, we performed Southern and Northern blot hybridizations. As shown in Figure 1c and d, whereas CapG mRNA was not detected in RBT cells, no gross structural alterations of the *CapG* gene were observed in any of the RB cell lines. This suggested that CapG expression was repressed at the transcriptional level at the stage of progression from anchorage independence to tumorigenicity in the RB cell lineage.

Ectopic CapG expression suppressed RBT cell tumorigenicity

The finding that loss of CapG expression was associated with progression from the anchorage-independent to the tumorigenic stage of the RB cell lineage led us to hypothesize that the CapG protein played a role in tumor suppression. To test this possibility, we introduced CapG cDNA into RBT cells by infection with a recombinant retroviral vector, and isolated cell clones expressing CapG protein. Levels of CapG protein in these clones were largely comparable to those observed

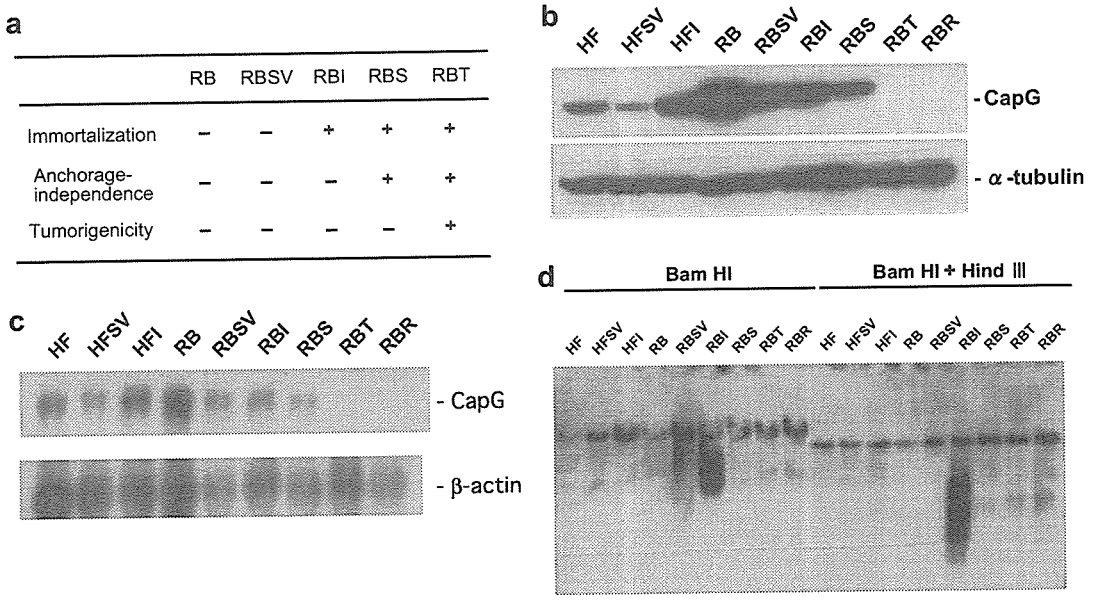


Figure 1 Loss of CapG expression in a tumorigenic cell line of the RB cell lineage. (**a**) Multistep transformation of the RB cell lineage. (**b**) CapG protein expression in the RB cell lineage detected by Western blot analysis. α-Tubulin was used as a loading control. RBR is a cell line reconstituted from an RBT tumor from a nude mouse. HF: human diploid fibroblasts; HFSV: HF cell clone transfected with SV40 early genes; HFI: immortalized HF cell line obtained from HFSV. (**c**) CapG mRNA expression in the RB cell lineage analysed by Northern blot. The RNA blot was hybridized with ^{32}P-labeled CapG cDNA (1.2 kb). β-Actin was used as a loading control. (**d**) No structural alteration of the *CapG* gene were observed in the RB cell lineage cell lines. Genomic DNA was analysed by Southern blot hybridization using ^{32}P-labeled CapG cDNA as a probe.

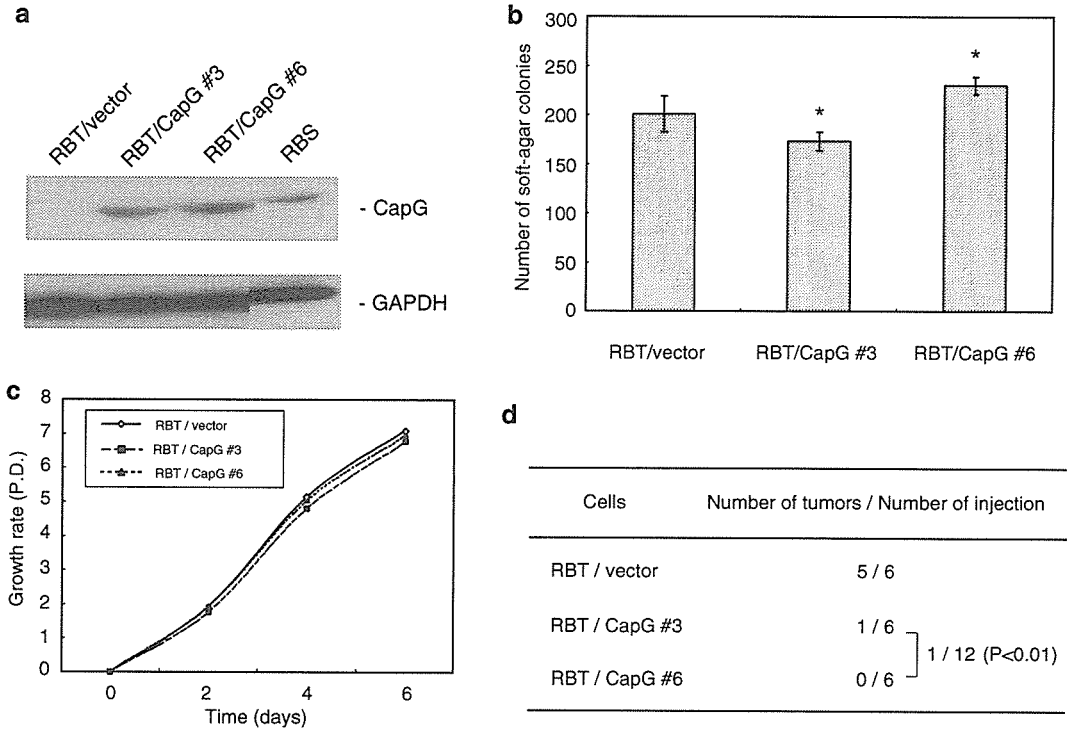


Figure 2 Tumorigenicity and growth properties of the RBT cell line infected with a CapG retroviral construct. (**a**) CapG protein expression as determined by Western blot. RBT/vector and RBT/CapG are mock- and CapG retrovirus-infected RBT cell clones, respectively. RBS is an anchorage-independent but non-tumorigenic cell line of the RB cell lineage. GAPDH was used as a loading control. (**b**) Anchorage-independent growth of RBT cell clones expressing CapG protein. Ten thousand cells were plated in 0.35% soft agar medium, and colonies larger than 0.125 mm counted 2 weeks later. The experiment was performed in triplicate. Error bars represent standard deviation from mean value. Asterisk shows 'nonsignificance' ($P > 0.05$). (**c**) Growth of RBT cell clones expressing CapG protein. Twenty-five thousand cells were seeded into liquid medium in 24-well culture plates. At the times indicated, cells in triplicate wells were trypsinized and counted using a hemocytometer. (**d**) Tumorigenicity in nude mice of RBT cell clones expressing CapG protein. Ten million cells were injected subcutaneously into nude mice. Data represent number of mice with tumors at 7 weeks after injection/number of mice injected.

in RBS cells, which grow anchorage independently but do not form tumors in nude mice (Figures 1a and 2a, and data not shown). The growth rate in liquid media and colony-forming ability in soft agar of RBT cell clones expressing CapG protein were similar to those of parental non-infected RBT cells (Figure 2b and c), which indicated that proliferation and anchorage-independent growth were not affected by CapG expression. However, when injected subcutaneously into nude mice, the tumor-forming capacity of CapG-expressing RBT cells was greatly reduced compared to non-infected RBT cells and RBT cells infected with a control retroviral vector (Figure 2d). These observations indicated that *CapG* gene suppressed RBT tumorigenicity, and that tumor suppression was independent of growth in liquid culture or soft agar.

Lack of CapG expression in various human cancer cell lines

Although ectopic CapG expression suppressed RBT cell tumorigenesis, the RB cell lineage is a series of artificially transformed cell lines rather than actual cancer cell lines. Therefore, we examined whether CapG expression was altered in human cancer cells. As shown in Figure 3a, CapG mRNA was ubiquitously expressed in normal human tissues, although expression levels were not uniform. However, some cancer cell lines, especially those derived from human stomach cancer (1/3), cutaneous melanoma (1/3), small-cell lung cancer (3/5) and lung adenocarcinoma (1/5) had completely lost CapG expression (Figure 3b). Downregulated or reduced CapG protein expression was also observed at a high frequency in various lung cancers (Figure 3b). These results suggested that loss of CapG expression may contribute to the development of human cancers.

CapG suppressed the tumorigenicity of stomach cancer cell line AZ521

To examine whether ectopic CapG expression could also suppress the tumorigenicity of an established human cancer cell line, we introduced CapG cDNA into the human stomach cancer cell line AZ521 that lacked CapG protein expression as shown in Figure 3 and isolated three cell clones that expressed CapG protein (Figure 4a). These CapG-expressing AZ521 clones exhibited similar growth rates and colony-forming ability compared to control AZ521 cells (Figure 4b and c), which again suggested that CapG protein

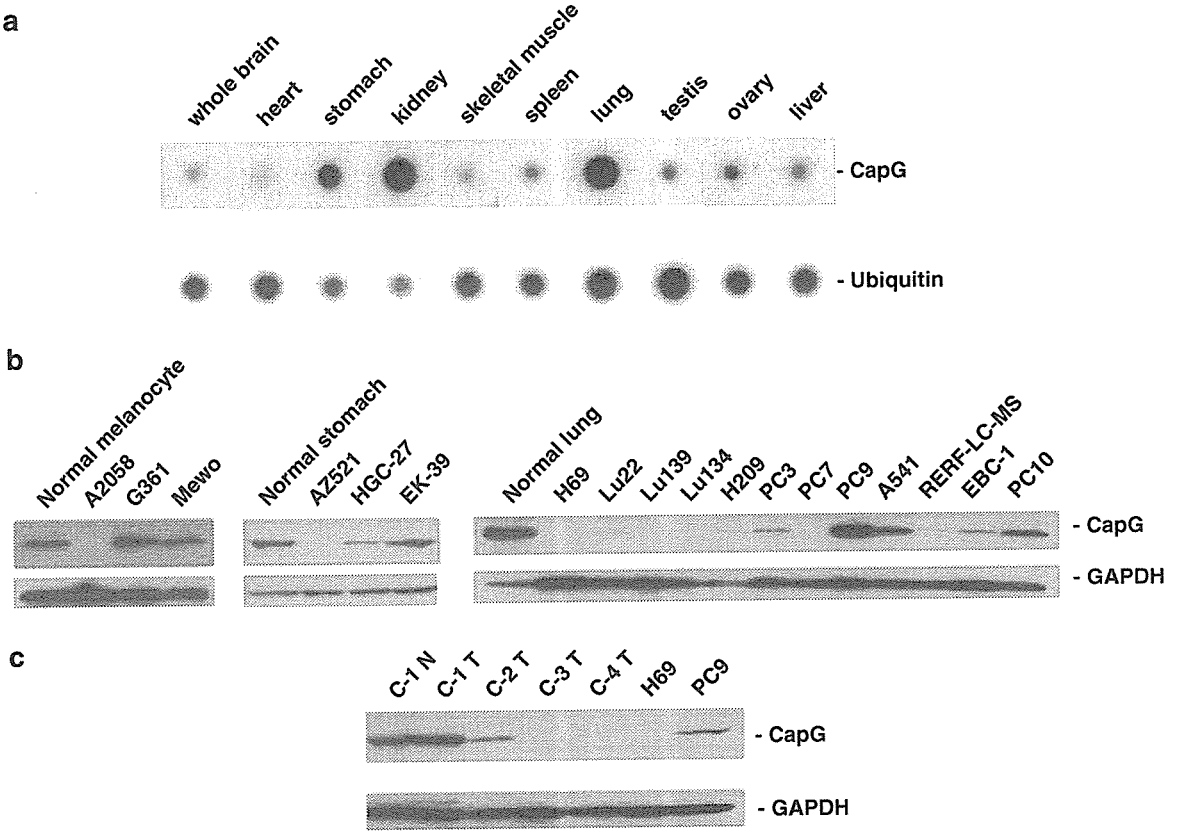


Figure 3 CapG expression in human normal tissues, cancer cell lines and primary tumors. (**a**) Expression of CapG mRNA in normal human tissues. Human RNA blot (Clonthech) was subjected to Northern blot hybridization with ^{32}P-labeled CapG cDNA. Several major tissues were selected from the original autoradiogram (Supplementary Figure 4). Ubiquitin cDNA supplied by the manufacturer was used as an internal control. (**b**) Expression of CapG protein in human melanoma (A2058, G361, MeWo), stomach cancer (AZ521, HGC-27, EK-39), small-cell lung cancer (H69, Lu22, Lu139, Lu134, H209), lung adenocarcinoma (PC3, PC9, PC9, A541, RERF-LC-MS) and lung squamous-cell carcinoma (EBC-1, PC10) cell lines. Expression levels were determined by Western blot. GAPDH was used as an internal control. (**c**) Expression of CapG protein in human small-cell lung cancer tissues. Lysates from four tumor tissues (C-1T to C-4T) and one normal lung tissue (C-1N) were subjected to Western blot. GAPDH was used as a loading control.

expression did not affect cell growth in liquid media or soft agar. However, CapG-expressing AZ521 clones did not produce tumors in nude mice (Figure 4d), which suggested that CapG expression suppressed the tumorigenicity of the human cancer cells in nude mice.

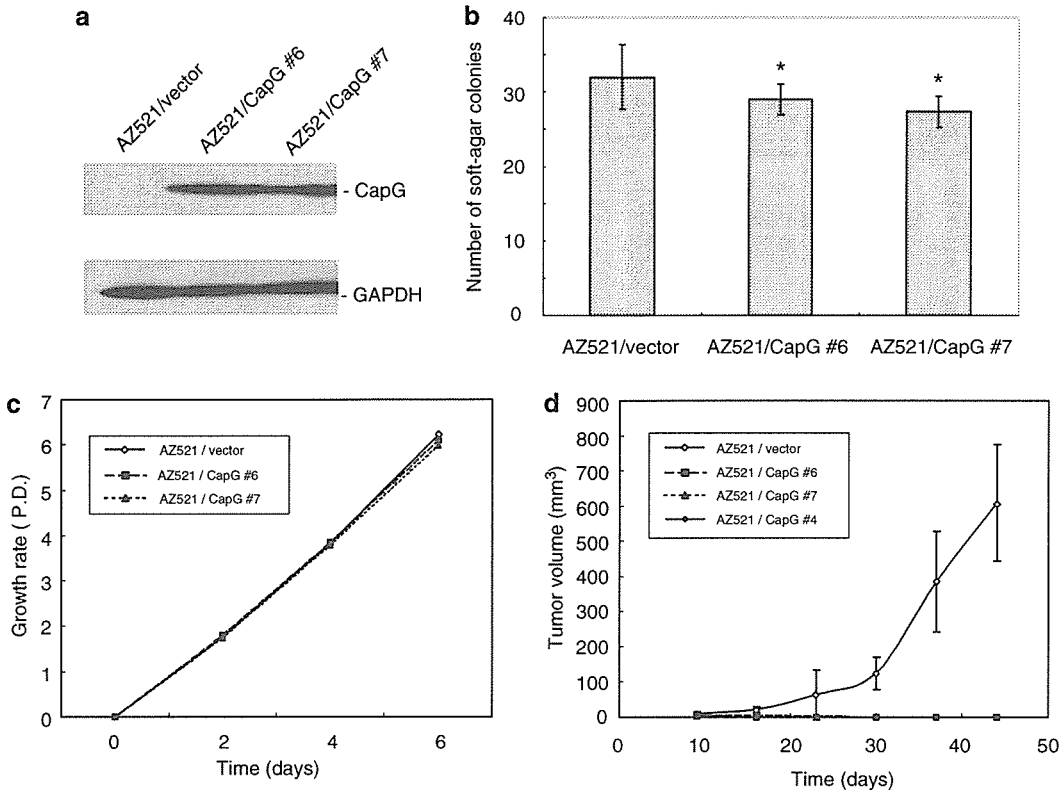
Loss of CapG protein expression was also observed in primary tumors

As most cancer cell lines are cultured *in vitro* over long periods of time, it can be argued that cancer cell lines do not necessarily accurately reflect primary tumor cells. To examine whether primary tumors also showed loss of CapG protein expression, we performed Western blot analysis using human tumor tissues isolated from cases of stomach cancer (four cases), small-cell lung cancer (four cases) and melanoma (two cases). The results showed that although downregulation of CapG protein was not observed in the stomach cancer and melanoma samples, CapG protein levels were greatly reduced in three of the four small-cell lung cancer samples (Figure 3c). Although some CapG protein was still present, tumor tissues can also contain normal cells such as stromal cells or infiltrating lymphocytes that would express normal levels of CapG. Therefore, our results strongly suggested that the development of certain human primary cancers could also involve the loss of the *CapG* gene function.

Taken together, our findings indicated that CapG may act as a tumor suppressor protein involved in tumorigenicity in nude mice but not in cell growth in liquid media or soft agar.

Effect of ectopic CapG expression on microfilament organization and cell motility

Dynamic reorganization of the actin cytoskeleton is an underlying factor in the tumor-related processes of invasion and metastasis. Because the CapG protein is known to possess actin-modulating activity (Silacci *et al*., 2004), it is possible that an alteration of microfilaments contributes to the acquisition or loss of tumorigenicity. Therefore, microfilament organization was examined in the RB cell lineage. Although the expression level of actin protein was comparable in all of the RB cell lines regardless of the presence or absence of



e

Cells	Number of tumors / Number of injection	
AZ521 / vector	3 / 3	
AZ521 / CapG #4	0 / 3	0 / 2 (P<0.01)
AZ521 / CapG #6	0 / 3	
AZ521 / CapG #7	0 / 3	

Figure 4 Suppression of tumorigenicity of a stomach cancer cell line AZ521 infected with CapG retroviral vector. (**a**) CapG protein expression. Cell lysates from AZ521 cell clones infected with control or CapG retroviral constructs were subjected to Western blotting with GAPDH used as a loading control. (**b**) Anchorage-independent growth of AZ521 cells expressing CapG protein. Ten thousand cells were plated in 0.35% soft agar medium, and the number of colonies larger than 0.125 mm counted 2 weeks later. The experiment was performed with triplicate samples. The data represent the mean value and standard deviation. Asterisk shows 'nonsignificance' ($P > 0.05$). (**c**) Growth of AZ521 cells expressing the CapG protein. Twenty-five thousand cells were seeded into liquid medium in 24-well culture plates. At the times indicated, cells in triplicate wells were trypsinized and counted using a hemocytometer. (**d**) Suppression of tumor formation by ectopic CapG expression. AZ521 cell clones (2.5×10^6) were injected subcutaneously into nude mice, and the size of tumors that appeared measured every week. Data represent the mean value ($n = 3$ each) and standard deviation. (**e**) Tumorigenicity of AZ521 cell clones expressing CapG protein. Data were obtained from the experiment (**d**), and represent number of mice with tumors at 7 weeks after injection/number of mice injected.

CapG (Figure 5a), actin filament organization was greatly altered with malignant progression of the RB cell lineage (Figure 5b). Microfilaments were well organized and stretched long in diploid RB cells, and were maintained, though much finer and shorter, even in the anchorage-independent cell line RBS (Figure 5b). In contrast, the actin stress fibers seemed to disappear in the tumorigenic RBT cell line, but then reassembled to form a thick bundle at the perinuclear region (Figure 5b). These findings led us to hypothesize that the tumorigenicity of the RBT cell line might be caused, at least in part, by alteration of microfilament organization. One caveat, however, is that the ectopic CapG expression that resulted in the suppression of tumorigenicity in RBT cells did not affect formation of the actin bundle (Figure 5b).

We next examined the effect of ectopic CapG expression on cell motility that is believed to have some relation to actin cytoskeleton. The same RBT transfectants that were used for the tumorigenicity test and microfilament staining were subjected to a migration assay. As shown in Figure 5c, ectopic expression of CapG enhanced migration activity of the RBT cell line. This finding implies that cell motility and tumorigenicity may be controlled by different functions of the CapG protein.

These results suggest that the actin-modulation function of CapG may not be involved in the

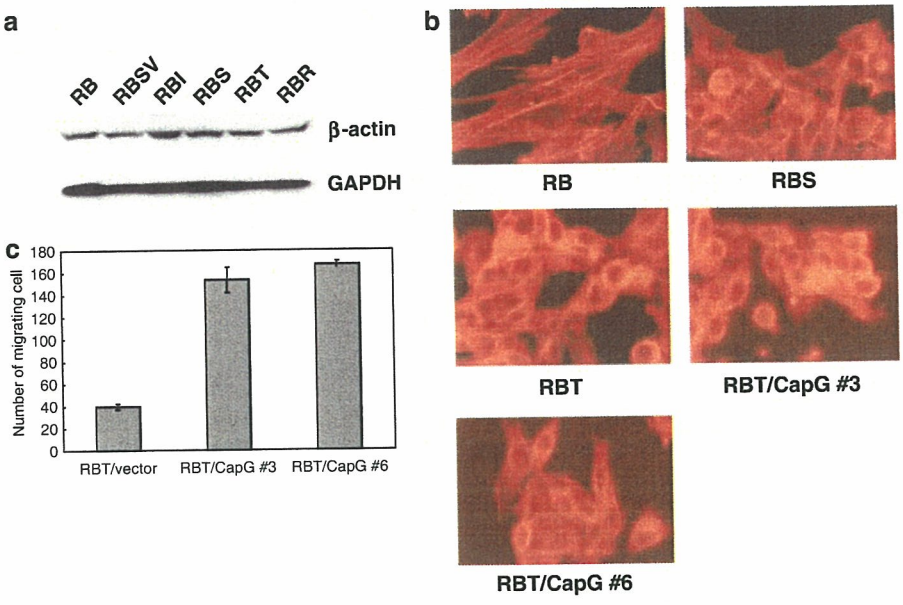


Figure 5 Effect of ectopic CapG expression on microfilament organization and cell motility. (**a**) Expression level of β-actin protein. Total cell lysates of the RB cell lines were subjected to Western blot analysis with anti-β-actin and anti-GAPDH antibodies. (**b**) Cytoskeleton organization of the RB cell lines. The RB cell lines were fixed with 2% paraformaldehyde, and stained with BODIPY 558/568 phalloidin (Molecular Probes, Eugene, OR, USA). (**c**) Enhanced migration of the RBT cells by overexpression of CapG protein. The RBT cell line was transfected with CapG expression vector or empty vector, and stably transfected clones were isolated. These cell clones (4×10^4 cells) were subjected to cell migration assay. Data represent mean value ($n = 3$ each) and standard deviation, and are significant ($P < 0.01$).

suppression of tumorigenicity. Clarification of the mechanism of tumor suppression by CapG will necessitate the isolation and characterization of CapG-interacting proteins.

Discussion

In this paper, we analysed the RB cell lineage, a human model of multistep tumorigenesis, to identify genes involved in malignant progression. As the different RB cell lines were derived from a single line of human diploid fibroblasts, it is likely that they shared a uniform genetic background with alterations limited to the expression of genes associated with malignant progression. Therefore, the RB cell lineage represented an effective tool for the identification of oncogenes or tumor suppressor genes responsible for each step of malignant progression. The cell lineage was not transfected with exogenous genes for malignant progression from immortalized RBI cell lines, such that the malignant progression observed in this cell lineage was likely to be due, at least in part, to the inactivation of endogenous tumor suppressor genes (Oka *et al*., 1999). In this way, the process of malignant progression in the RB cell lineage may mimic the process that occurs in naturally occurring cancer. Therefore, identification of genes with altered expression in response to malignant progression using the RB cell lineage may be a useful tool to help elucidate the mechanisms of the multistep tumorigenic process.

By comparing gene expression profiles of the RB cell lineage, we found that CapG expression was lost at the transcriptional level at the stage of progression from the non-tumorigenic to the tumorigenic state. Similar downregulation of CapG expression was also observed in various human cancer cell lines and cancer tissues. These findings allowed us to hypothesize that the CapG protein may be involved in the tumorigenic progression of cancer cells as a tumor suppressor gene. To examine this possibility, we introduced CapG cDNA into the tumorigenic line of the RB cell lineage, RBT, as well as the gastric cancer cell line AZ521, both of which showed no endogenous CapG protein expression. We then tested the ability of these cell lines to form tumors in nude mice and to form colonies in soft agar. Our results indicated that ectopic *CapG* gene expression suppressed tumorigenicity, but did not affect the anchorage-independent growth of these cell lines.

CapG, also known as gCap39, Mbh1 or MCP, is a 348-amino-acid protein that is ubiquitously expressed in normal tissues, being particularly abundant in macrophages (Yu *et al*., 1990; Prendergast and Ziff, 1991; Dabiri *et al*., 1992). CapG is a member of the gelsolin family of actin filament modulating proteins that also includes gelsolin, villin, adseverin, advellin, supervillin and flightless I (Silacci *et al*., 2004). However, CapG has features that distinguish it from other gelsolin family proteins. CapG has only three repeated gelsolin-like domains, in contrast to the usual six domains present in other gelsolin family proteins, and it lacks the actin-severing activity exhibited by the other family members. Another unique characteristic of CapG is its subcellular

localization. Whereas CapG localizes to both the cytoplasm and nucleus, the other gelsolin family members are present only in the cytoplasm. Therefore, CapG may have a function in addition to actin or cytoskeleton modulation. Indeed, it has been reported that CapG represses transcriptional activation (De Corte *et al.*, 2004), although it was not clear which gene was trans-repressed by CapG.

Among these CapG properties, the cytoskeleton-modulating function may be dispensable for the suppression of tumorigenicity, as ectopic CapG expression in the tumorigenic RBT cell line did not affect cytoskeletal appearance (Figure 5b) but did suppress tumorigenicity. CapG expression and nuclear localization patterns were also unchanged between the non-tumorigenic cell line RBS and the non-tumorigenic transfectants, RBT/CapG cells and AZ521/CapG cells, which expressed ectopic CapG protein (Supplementary Figure 5). This persistent expression suggests that tumor suppression does not result from abnormal expression or localization of CapG protein.

On the other hand, CapG reportedly possesses an oncogenic function involved in the control of cell migration or invasion. When cells were transfected with CapG expression vector, they acquired stimulated migration/invasion activity (Pellieux *et al.*, 2003; De Corte *et al.*, 2004). Enhanced migration of RBT cells was also observed in this study after ectopic expression of CapG (Figure 5c). This latter experiment utilized the same transfected cell clones that were examined for tumorigenicity (Figures 2d and 5c). Consequently, the indication is that CapG overexpression leads to the simultaneous suppression of tumorigenicity and activation of migration/invasion. Thus, the implication is that activated migration/invasion does not hamper tumor suppression. In this regard, it is noteworthy that CapG protein overexpression in ocular melanoma and glioblastoma has been reported (Van Ginkel *et al.*, 1998; Lal *et al.*, 1999). Although the effect of the overexpression on tumorigenesis has not been clarified, it may be interesting to examine whether these cancers contain cells with *CapG* gene mutations.

Given that CapG function can be activated via the Ras/MAPK signaling pathway (De Corte *et al.*, 2004), it is possible that there may be a relationship between loss of CapG expression and activation of the Ras/MAPK signaling pathway. It is well known that the *K-Ras*, *N-Ras* and *BRAF* genes are frequently mutated and activated in melanoma and stomach cancer (Davies *et al.*, 2002; Rajagopalan *et al.*, 2002). Furthermore, our results showed that some melanoma and stomach cancer cell lines had lost CapG protein expression, which suggests that Ras/MAPK mutations may affect CapG protein expression levels. Therefore, we examined the melanoma and stomach cancer cell lines used in this paper for the presence of K-Ras, N-Ras and BRAF mutations by RT–PCR and sequencing. However, we found no relationship between mutation of these genes and CapG protein levels (data not shown). Further study is necessary to clarify the mechanism of tumor suppression by CapG.

Gelsolin, the prototype gelsolin family protein, also exhibits tumor suppressor activity in certain human cancer cell lines. Downregulation of gelsolin has been observed at a high frequency in various cancer cell lines, and ectopic expression of gelsolin suppressed tumorigenicity of bladder and lung cancer cell lines (Tanaka *et al.*, 1995; Sagawa *et al.*, 2003). However, in these cases, gelsolin suppressed not only tumorigenicity but also anchorage-independent cell growth, which suggests that the mechanisms of tumor suppression by CapG and gelsolin are different.

In this study, based on RB cell lineage cell lines, we attempted to investigate the mechanisms that underlie multistep malignant progression and identified a candidate tumor suppressor gene, *CapG*. Loss of CapG protein and mRNA expression was observed at the progression from the non-tumorigenic to tumorigenic state, and ectopic CapG expression in tumorigenic RBT cells resulted in the inhibition of tumorigenicity, but not anchorage-independent cell growth. CapG protein levels in the RB cell lineage coincided well with the activity of the *CapG* gene, which indicated that our cell lines were a useful model to clarify the mechanisms of malignant progression. Although the molecular mechanism of CapG-induced tumor suppression remains unclear, we expect our model cell lines will contribute greatly to the understanding of tumorigenesis.

Materials and methods

Cell culture

RB cell lines were cultured in Dulbecco's modified minimal essential medium (DMEM) supplemented with 10% fetal bovine serum (FBS) as described previously (Oka *et al.*, 1999). The gastric cancer cell line AZ521 was cultured in minimal essential medium (MEM) supplemented with 10% FBS. Cancer cell lines were obtained from the Japanese Cancer Research Resources Bank, Health Science Research Resources Bank or Riken Cell Bank.

Western blot analysis

Cells were lyzed with RIPA buffer (0.15 M NaCl, 50 mM Tris-HCl (pH 7.4), 1 mM ethylenediaminetetraacetic acid, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 100 μg/ml phenylmethylsulfonyl fluoride and 0.25 TIU/ml aprotinin). The cell lysates were run on 5–20% SDS–polyacrylamide gel and electroblotted onto polyvinylidene difluoride membrane. The membranes were incubated successively with rabbit polyclonal anti-CapG (TransGenic, Kumamoto, Japan), mouse monoclonal anti-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Chemicon International, Temecula, CA, USA) or α-tubulin (Sigma, St Louis, MO, USA) antibody and then with horseradish peroxidase-conjugated anti-rabbit or mouse IgG antibody (Cell Signaling Technology, Beverly, MA, USA). Protein bands were detected using enhanced chemiluminescence reagent (Amersham, Piscataway, NJ, USA).

Southern blot hybridization

DNA was digested with restriction enzymes, subjected to 1% agarose gel electrophoresis and transferred onto nylon

membranes. Membranes were then hybridized with ^{32}P-labeled probe, washed and autoradiographed.

Northern blot hybridization
RNA isolated using an SV total RNA isolation kit (Promega, Madison, WI, USA) was subjected to formaldehyde-1.2% agarose gel electrophoresis, blotted onto nylon membranes, hybridized with ^{32}P-labeled probe, washed and autoradiographed.

Construction of retrovirus vector and virus production
pCX4bsr, a Moloney murine leukemia virus-based retrovirus vector containing the blasticidin S resistance gene (*bsr*) as a selectable marker, was constructed by Akagi *et al.* (2000). pCX4-CapG was constructed by inserting full-length human CapG cDNA into the multi-cloning site of the pCX4bsr vector.

Phoenix-A cells were transfected with pCX4bsr or pCX4-CapG constructs using FuGene 6 (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's protocols. Two days after transfection, the culture supernatants were collected and stored at −70°C until use.

Tumorigenicity in nude mice
Female 4-week-old BALB/cJ (nu/nu) mice were injected subcutaneously in the back with CapG- or empty virus-infected cells in 0.2 ml of DMEM (without serum). After 7 weeks, tumor size was measured in two dimensions using hand calipers, and tumor volume calculated by the formula $0.5 \times L \times W^2$, where L and W are the length and width of a tumor, respectively (Sagawa *et al*, 2003).

Anchorage-independent growth assay
Anchorage-independent growth was assessed by colony-forming ability in soft agar. Ten thousand cells were inoculated into 0.35% agarose in DMEM supplemented with 10% FBS per 60 mm dish. After 2 weeks' incubation, the number of colonies (> 0.125 mm in diameter) was scored.

Cell migration assay
Cells (4×10^4) were suspended in 200 μl of serum-free DMEM containing 0.1% bovine serum albumin, and plated in upper chamber of Chemotaxicell (Kurabo, Osaka, Japan). The lower chamber contained 600 μl of phosphate-buffered saline supplemented with 6 μg/ml fibronectin. After incubation for 6 h at 37°C, cells were fixed with 80% methanol, and stained with hematoxylin. The number of cells migrated through membrane was counted under microscope.

Acknowledgements

We thank Dr Akagi for providing the pCX4 retrovirus vector. SH is a recipient of a grant from the Charitable Trust Osaka Cancer Researcher-Fund.

References

Akagi T, Shishido T, Murata K, Hanafusa H. (2000). *Proc Natl Acad Sci USA* **97**: 7290–7295.
Dabiri GA, Young CL, Rosenbloom J, Southwick FS. (1992). *J Biol Chem* **267**: 16545–16552.
Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S *et al* (2002). *Nature* **417**: 949–954.
De Corte V, Van Impe K, Bruyneel E, Boucherie C, Mareel M, Vandekerckhove J *et al.* (2004). *J Cell Sci* **117**: 5283–5292.
Hahn WC, Weinberg RA. (2002). *Nat Rev Cancer* **2**: 331–341.
Lal A, Lash AE, Altschul SF, Velculescu V, Zhang L, McLendon RE *et al.* (1999). *Cancer Res* **59**: 5403–5407.
Oka K, Tomonaga Y, Nakazawa T, Ge H-Y, Bengtsson U, Stanbridge EJ *et al.* (1999). *Genes, Chromosome Cancer* **26**: 47–53.
Pellieux C, Desgeorges A, Pigeon CH, Chambaz C, Yin H, Hayoz D *et al.* (2003). *J Biol Chem* **278**: 29136–29144.
Prendergast GC, Ziff EB. (1991). *EMBO J* **10**: 757–766.
Rajagopalan H, Bardelli A, Lengauer C, Kinzler KW, Vogelstein B, Velculescu VE. (2002). *Nature* **418**: 934.
Rajogopalan H, Nowak MA, Vogelstein B, Lengauer C. (2003). *Nat Rev Cancer* **3**: 695–701.
Sagawa N, Fujita H, Banno Y, Nozawa Y, Katoh H, Kuzumaki N. (2003). *Br J Cancer* **88**: 606–612.
Silacci P, Mazzolai L, Gauci C, Stergiopulos N, Yin HL, Hayoz D. (2004). *Cell Mol Life Sci* **61**: 2614–2623.
Tanaka M, Mullauer L, Ogiso Y, Fujita H, Moriya S, Furuuchi K *et al.* (1995). *Cancer Res* **55**: 3228–3232.
Van Ginkel PR, Gee RL, Walker TM, Hu D-N, Heizmann CW, Polans AS. (1998). *Biochim Biophys Acta* **1448**: 290–297.
Yu F-X, Johnston PA, Sudhof TC, Yin HL. (1990). *Science* **250**: 1413–1415.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc).

RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2006; **20**: 2589–2595
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rcm.2595

'Information-Based-Acquisition' (IBA) technique with an ion-trap/time-of-flight mass spectrometer for high-throughput and reliable protein profiling

Toshiyuki Yokosuka[1]*, Kiyomi Yoshinari[1], Kinya Kobayashi[1], Atsushi Ohtake[1], Atsumu Hirabayashi[2], Yuichiro Hashimoto[2], Izumi Waki[2] and Toshifumi Takao[3]

[1]Hitachi Research Laboratory, Hitachi, Ltd., Omika 7-1-1, Hitachi, Ibaraki 319-1292, Japan
[2]Central Research Laboratory, Hitachi, Ltd., Higashi-Koigakubo 1-280, Kokubunji, Tokyo 185-8601, Japan
[3]Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565-0871, Japan

Received 26 December 2005; Revised 1 June 2006; Accepted 2 June 2006

Highly complex protein mixtures can be analyzed after proteolysis using liquid chromatography/mass spectrometry (LC/MS). In an LC/MS run, intense peptide ions originating from high-abundance proteins are preferentially analyzed using tandem mass spectrometry (MS^2), so obtaining the MS^2 spectra of peptide ions from low-abundance proteins is difficult even if such ions are detected. Furthermore, the MS^2 spectra may produce insufficient information to identify the peptides or proteins. To solve these problems, we have developed a real-time optimization technique for MS^2, called the Information-Based-Acquisition (IBA) system. In a preliminary LC/MS run, a few of the most intense ions detected in every MS spectrum are selected as precursors for MS^2 and their masses, charge states and retention times are automatically registered in an internal database. In the next run, a sample similar to that used in the first run is analyzed using database searching. Then, the ions registered in the database are excluded from the precursor ion selection to avoid duplicate MS^2 analyses. Furthermore, real-time *de novo* sequencing is performed just after obtaining the MS^2 spectrum, and an MS^3 spectrum is obtained for accurate peptide identification when the number of interpreted amino acids in the MS^2 spectrum is less than five. We applied the IBA system to a yeast cell lysate which is a typical crude sample, using a nanoLC/ion-trap time-of flight (IT/TOF) mass spectrometer, repeating the same LC/MS run five times. The obtained MS^2 and MS^3 spectra were analyzed by applying the Mascot® (Matrix Science, Boston, MA, USA) search engine to identify proteins from the sequence database. The total number of identified proteins in five LC/MS runs was three times higher than that in the first run and the ion scores for peptide identification also significantly increased, by about 70%, when the MS^3 spectra were used, combined with the MS^2 spectra, before being subjected to Mascot® analysis. Copyright © 2006 John Wiley & Sons, Ltd.

Highly complex protein mixtures can be analyzed after proteolysis using liquid chromatography/mass spectrometry (LC/MS). Recently, special attention has been paid to the analysis of the low-abundance proteins that are thought to relate to diseases. However, in an LC/MS run, the peptide ions originating from high-abundance proteins are preferentially analyzed by tandem mass spectrometry (MS^2). It is difficult to obtain MS^2 spectra of peptide ions from low-abundant proteins even if the ions are detected. This is because only a limited number of ions are selected as precursors of the MS^2 analyses. In general, it is necessary to separate the crude sample to reduce the influence of the high-abundance proteins using gel electrophoresis or multi-dimensional LC/MS thus spending a good deal of time in complex preprocessing. Furthermore, there is the problem that the analysis of the further low-abundance proteins will become difficult using the current method when the sensitivity of mass spectrometer improves. Therefore, in order to increase the throughput of the analysis of low-abundance proteins, it is necessary to reduce the number of LC/MS runs, and ions detected in MS, which were previously selected as MS^2 precursor ions, should be excluded from subsequent MS^2 precursor ion candidates.

So far, a number of different approaches have been employed to identify proteins from the MS^2 spectra. The most common approaches rely on available databases to match the experimental MS^2 spectra.[1,2] However, some MS^2 spectra may have insufficient information to identify the peptides or proteins, as in the reported cases of neutral loss or non-mobile protons.[3–9] MS^2 data do not always provide

*Correspondence to: T. Yokosuka, Hitachi Research Laboratory, Hitachi, Ltd., Omika 7-1-1, Hitachi, Ibaraki 319-1292, Japan.
E-mail: toshiyuki.yokosuka.gc@hitachi.com
Contract/grant sponsors: Hitachi High Technologies, Ltd; New Energy and Industrial Technology Development Organization (NEDO).

RCM

sufficient fragment ion peaks for identifying peptides with a high accuracy. In such cases, MS^3 analysis, which is a technique that could obtain more information about fragment species of a precursor ion, will be efficient for identifying as many peptides in a mixture as possible. However, MS^3 analysis usually requires more time than MS^2 analysis. Therefore, unnecessary MS^3 analysis reduces the total throughput of the MS analysis. So, it is necessary to judge whether the MS^2 spectrum have sufficient information to identify the peptides during the measurement, and MS^3 analysis should be performed only when it is needed.

In this paper we describe a new approach using software, called Information-Based-Acquisition (IBA), in which the sequence of tandem mass spectrometry is automatically optimized, so that precursor ions should be selected to increase identified proteins with high accuracy.

INFORMATION-BASED ACQUISITION (IBA)

Figure 1 shows the workflow for protein profiling with an ion trap time-of-flight (IT/TOF) mass spectrometer controlled by the IBA system. The IBA, which is a real time-optimization technique for tandem mass spectrometry (MS^2), consists of the following two main functions: (1) precursor ion selection with an internal database for MS data, and (2) automated MS^3 analysis with real-time *de novo* sequencing[10–13] for MS^2 spectra. To perform real-time optimization, it is preferable that the processing time for optimization is within 10 ms which is negligible enough for actual MS measurement. The features of each function are as follows.

Precursor ion selection with an internal database (IBA (I))

IBA (I) judges whether an ion species has already been analyzed in order to avoid duplicate analysis. Figure 2 shows the processing flow of IBA (I). The approach taken in IBA (I) can be summarized into three steps: (i) preprocessing, (ii) database searching, and (iii) scoring and selecting precursor ions.

(i) The first step consists of preprocessing of the raw MS data. This involves a new method for noise filtering (1-1), peak centering, as well as deconvolution of the multiply charged species to a neutral mass value (1-2). This step is very important for the judgment of the information on the ions. Therefore, optimal preprocessing of the data is the most important step for precursor ion selection with internal database searching. To verify the judgment accuracy of preprocessing, we applied our algorithm to 278 peaks included in the same spectra extracted at random from the LC/MS data of bovine serum albumin (BSA) digested by trypsin and checked them by the visual evaluation. The judgment accuracy of ion information using IBA, such as monoisotopic mass value and charge state, is more than 90%.

(ii) The second step is database searching (1-3). IBA gives a score to an ion only when the data set (mass value (m), retention time (RT) and charge state (z)) of the ion does not agree with the data set of the internal database, or a fixed time after the ion begins to be detected, or until the integrated value of the analyzed ion intensity becomes a definite value. The data set of the MS^2 precursor ions is automatically stored in the internal database (DB) which can currently store 40 000 data points for MS^2 precursor ions. The storage capacity of the system can be increased if necessary. In particular, judgment of the integrated value of the analyzed ion's intensity is effective for the analysis of low-abundance species. Because the signal-to-noise (S/N) ratio of the MS^2 spectra of low-abundance species is very poor, it is preferable to measure over a long time period in order to improve the S/N ratio by summing of spectra.

(iii) The third step is scoring and selection of the next MS^2 precursor ion. The intensity of the ion is scored (1-4). If the

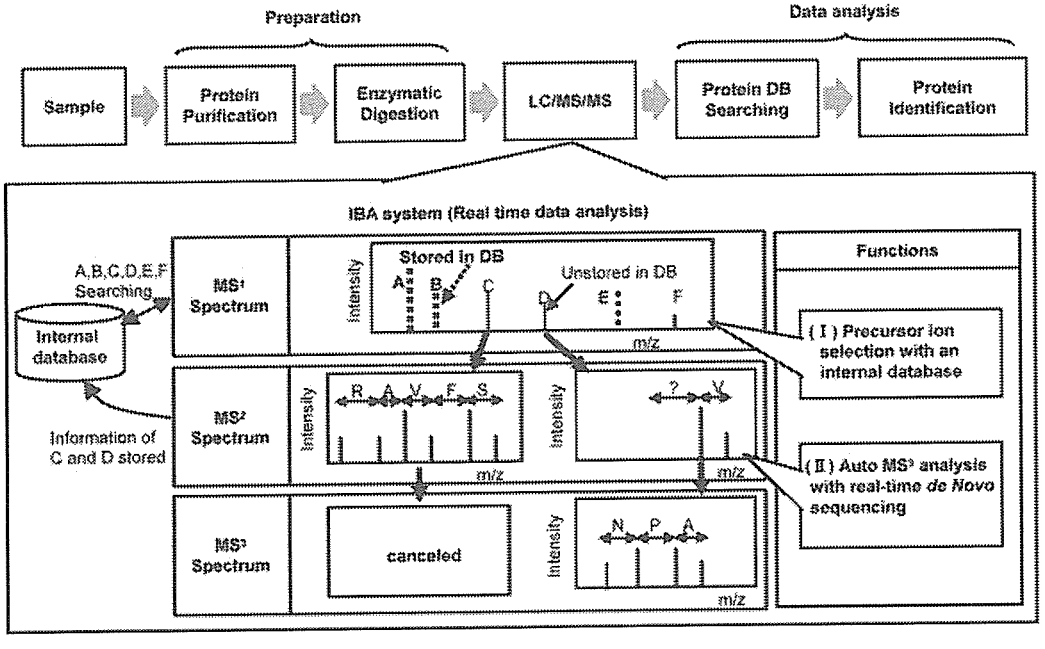


Figure 1. The workflow for protein profiling with an IT-TOF mass spectrometer controlled by an Information-Based-Acquisition (IBA) system.

DOI: 10.1002/rcm

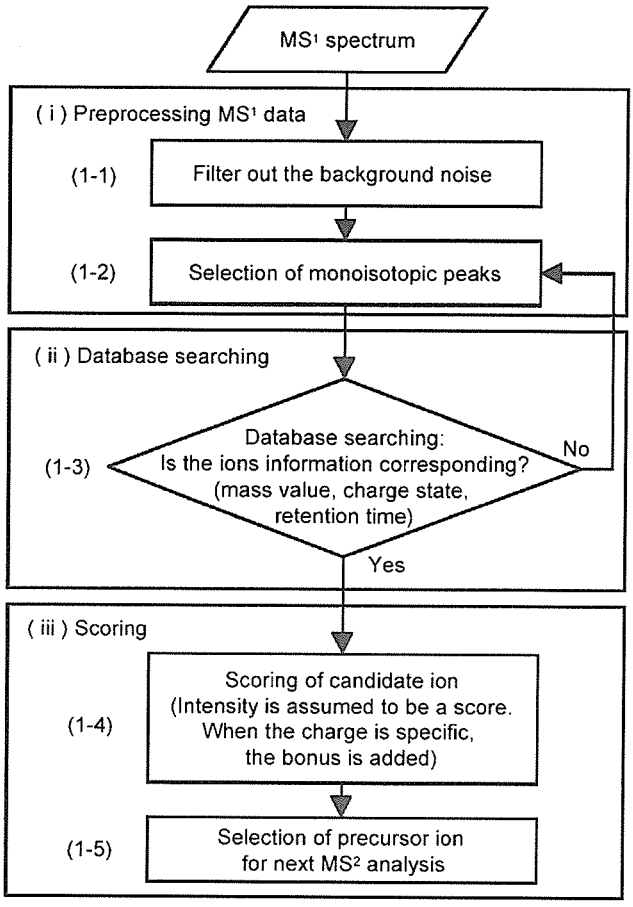


Figure 2. The processing flow of IBA (I).

user wants to give priority to a specific charge state, the parameter can be set beforehand. Finally, IBA selects the precursor ion for MS^2 analysis based on the score (1-5).

The processing time needed for the above-mentioned functions is less than 10 ms, which is the setup time of the next MS^2 acquisition. As reference data to judge whether or not the detected ions are identified with those previously selected as MS^2 precursor ions, using only mass value and charge state is insufficient. Retention time is also a very powerful reference data value to avoid repeating MS^2 on the same ion. Therefore, our system can refer to *m*, *z* and RT in a real-time manner so as to efficiently select low-abundance ions as MS^2 precursor ions. In standard techniques, such as Dynamic Exclusion™, MS^2 analysis of the same target is only allowed during a fixed time. In IBA (I), repetition of the same MS^2 analysis for the same precursor ion is allowed until the summation of intensity of the precursor ion becomes more than a definite value through all the analysis. As a result, a low-abundance ion can be analyzed microscopically more than a high-abundance ion.

We use the stored data of the DB after the Nth LC/MS run for the (N + 1)th LC/MS run. The processing time needed for this function is less than 10 ms, which is the setup time of the next MS^2 acquisition.

Auto MS^3 analysis (IBA (II))

When the MS^2 of a peptide gives insufficient fragment ion peaks, it might exhibit more fragmentation with MS^3. In order to judge the necessity of MS^3, it is essential to promptly figure out whether or not the current MS^2 data have sufficient information for identifying the peptide. The main feature of this step is to obtain the number of amino acid residues N_{aa} that can be interpreted from the MS^2 spectrum within 10 ms by concurrent assignment of b- and y-series ions to the observed peaks. The number of interpreted amino acid residues was previously considered as the criterion number N_c, which is used to judge whether the MS^2 data have sufficient information to identify the peptide. Based on this result, MS^3 could be conducted only when the number of the interpreted amino acids is below thr criterion number. In this study, we set the criterion number N_c to five.

Figure 3 shows the processing flow of IBA (II). The approach taken in IBA (II) can be broken down into four steps: (2-1) preprocessing, (2-2) candidate computation, (2-3) judgment of necessity of MS^3 analysis, and (2-4) selection of optimal precursor ion for MS^3 analysis.

(2-1) The first step is almost the same as the processing done with IBA (I). This involves a new method for noise filtering, peak centering, as well as deconvolution of the multiply charged species to singly charged ions for the raw MS^2 spectrum. This step is also very important for the interpretation of the MS^2 spectrum in IBA (II). The optimal preprocessing of data is an important step for the *de novo* sequencing using the MS^2 spectrum.

(2-2) The second step, candidate computation, is the critical step for predicting the amino acid sequences for a given computed precursor ion mass value. For this computation, a, b, c, x, y, z, the dehydration, and the deammonia peaks are considered. The basic assumption of our model is the evaluation of the intervals among all the peaks. For each interval value M, this new algorithm first computes the reward that a y (or b) ion has mass value M. If there is an interval close to M, the reward is equal to the logarithmic abundance of the peak, multiplied by a factor reflecting the co-existence of the derivative ion, such as x, z, y-18 (H_2O) and y-17 (NH_3) ions. The factor is used in the experimental detection probability when the dissociation technique is assumed to be collision-induced dissociation (CID) as reported by Fernandez-de-Cossio *et al.*[14] If the interval is not close to M, it is not used to predict the sequence. Next, the amino acid sequences are interpreted from the mass values of the precursor ions one by one by using only the peak with the corresponding mass value of the amino acid residues. This time, IBA (II) is simultaneously interpreted from the N- and C-termini to accelerate the processing.

(2-3) In the third step, the necessity of MS^3 analysis is judged from the number of interpreted amino acid residues (N_{aa}) with high scores from the predicted amino acid sequence. IBA (II) judges the necessity of MS^3 analysis from N_{aa}, which is a criterion parameter set by the user beforehand. If N_{aa} is greater than or equal to the criterion, IBA (II) judges that MS^3 analysis is not necessary and cancels the step ($N_{aa} \geq N_c$: case 1). In the case of angiotensin III (RVYIHPF, 1+), which gave sufficient peaks for sequencing, seven amino acid residues were predicted with a high score by IBA (II). In this case, the IBA will proceed to MS or MS^2 for other ion species. Meanwhile, if N_{aa} is less than the criterion, IBA (II) judges the MS^2 spectrum has insufficient information to accurately identify the correct peptide ($N_{aa} < N_c$: case 2). In

DOI: 10.1002/rcm

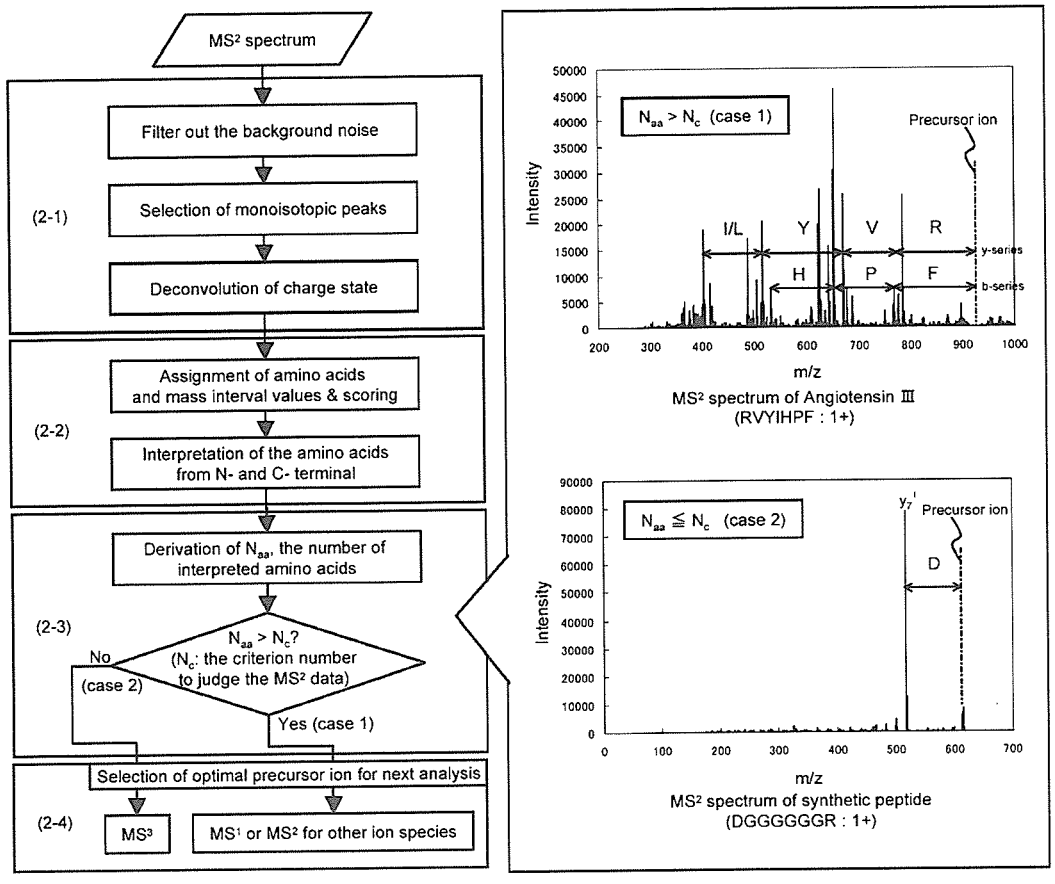


Figure 3. The processing flow of IBA (II).

the case of the synthetic peptide (DGGGGGGR, 1+), which gave insufficient peaks for sequencing, only asparatic acid (D) was predicted with a relevant score by IBA (II). In this case, the IBA will switch over to MS^3.

(2-4) Finally, if the N_{aa} is less than the criterion (case 2), IBA (II) automatically selects the precursor ion for MS^3 analysis. When the MS^3 analysis is executed, the y-ion is selected in this technique among the presumed peaks including the amino acid or the area where the score is low and uncertain by priority as a precursor ion of MS^3 analysis. This is because arginine (R) and lysine (K) of the basic amino acid residues that easily trap the proton ion added to the peptide are located in the C-terminus when the protein is digested by trypsin, and the possibility that the y-ion is detected with high intensity is high. Figure 4 shows the MS^3 spectrum of the y_7^+ ion of the synthetic peptide (DGGGGGGR) which was selected as a MS^3 precursor ion by IBA (II). By performing the MS^3 analyses to the y_7^+ ion, six amino acid residues (DGGGGG) of DGGGGGGR could be interpreted with high scores. Thus, some peptides will be identified by the auto MS^3 function.

Thus, IBA (II) is carried out on the obtained MS^2 spectrum and the MS^3 analysis is executed for accurate peptide identification when the number of interpreted amino acid residues in the MS^2 spectrum is less than the criterion which is set beforehand.

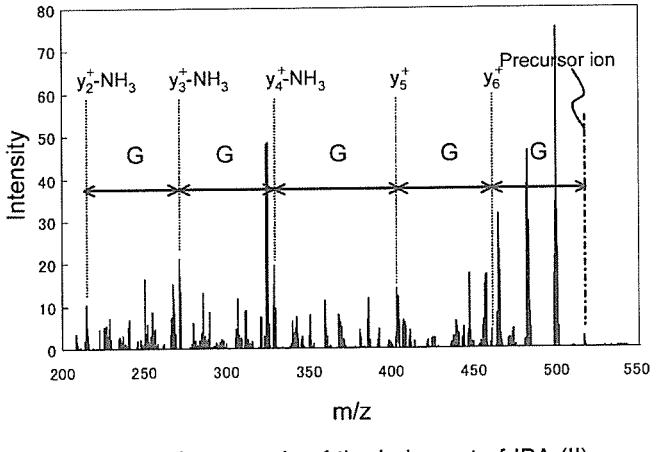


Figure 4. An example of the judgment of IBA (II).

EXPERIMENTAL

In order to achieve real-time optimization of tandem mass spectrometry, IBA was installed in a modified prototype model of a nanoLC/ion-trap/oaTOF mass spectrometer equipped with an electrospray ionization source (ESI) (Hitachi High-Technologies, Ltd., Tokyo, Japan).

We tested our technique by repeatedly analyzing a tryptic digest of soluble proteins extracted from *Saccharomyces cerevisiae* S288C, a typical crude sample. Since the nanoLC pump system is a non-split type, the retention time for detected ions is reproducible. The mobile phase A was water with 0.1% formic acid and B was a solution of 98% acetonitrile with 0.1% formic acid. A 180-min gradient was performed from 2% to 50% of mobile phase B at an effluent flow rate of 50 nL/min.

DOI: 10.1002/rcm

The sample was injected with an injection valve (M485, Upchurch Scientific, WA, USA) and was separated with a 600-mm-long monolithic column with a diameter of 30 mm (Kyoto Monotech, Kyoto, Japan). The backpressure of the nanoLC pump was about 5 MPa. A high voltage of 1.4 kV was applied to a fused-silica capillary spray chip with a diameter of 5 μm (New Objective, MA, USA) and the sampling orifice of the mass spectrometer was heated with a heater to 140°C.

In a preliminary LC/MS run, a few of the most intense ions detected in every MS spectrum are selected as precursors for MS^2, and their masses, charges, and retention times are automatically registered in an internal database. In the next run, the same sample as that used in the first run is analyzed using the database. The ions registered in the database are then excluded from the precursor-ion selection to avoid duplicating the analysis. With a repetition of such LC/MS runs, the number of ions registered in the database becomes significant, so low-abundance ions are readily analyzed. On the other hand, real-time *de novo* sequencing analysis is performed just after obtaining an MS^2 spectrum, and an MS^3 spectrum is obtained for accurate peptide identification only when the number of interpreted amino acids in the MS^2 spectrum is fewer than five.

Protein identification

The MS^2 identification was performed by protein database searching used by the Mascot® program[1] for all MS^2 and MS^3 spectra. The obtained MS^3 spectra were merged with the MS^2 spectra, which include precursor ions of MS^3 spectra. The database of *Saccharomyces cerevisiae* S288C (6332 sequences) was obtained from the Genome Information Broker (GIB), which is based on the data from the National Center for Biotechnology Information. Protein identifications are based on peptide identifications that have Mascot ion scores that correspond to at least a 95% probability of being correct.

RESULTS AND DISCUSSION

The information of the analyzed precursor that had been obtained from LC/MS analysis was automatically stored in an internal database, and used in the following LC/MS analysis.

Figure 5 indicates the time which was required to analyze each MS or MS^2 spectrum. The MS^1 spectrum required from 2 to 5 ms to optimize the next analysis. For the MS^2 spectrum, all processing was completed within about 2 ms. The processing time is within about 5 ms of the range of the retention time (from 30 to 100 min) when many ion species are detected. It is shown that on-line data-dependent MS^n ($n = 2, 3$) could be feasible and there is no negative influence on the throughput of the analysis.

Figures 6 and 7 show the effects of IBA (I). Figure 6 shows the distribution of the ionic species stored in an internal database. Almost 2000 kinds of different (mass value or charge state or retention time) ionic species were selected as MS^2 precursor ions and stored in an internal database in each LC/MS run. By repeating the same LC/MS run five times, 9078 kinds of different ion species were stored in total. Thus, by avoiding duplicate MS^2 analyses using IBA (I), the low-abundance peptides could be selected efficiently as MS^2 precursor ions.

Figure 7 shows the total number of identified proteins. To confirm the effect of a function of IBA (I), the result of the measurement without using an internal database is shown as auto MS/MS. The number of identified proteins increases almost linearly by a function of IBA (I), which selects different MS^2 precursor ions in each LC/MS run. In the first LC/MS run, 162 kinds of proteins were identified. On the other hand, 533 kinds of proteins were identified by repetition of the same LC/MS run five times. Thus, the number of proteins identified by running the analysis five times increases by 3.3 times the number of proteins identified in the first LC/MS run. On the other hand, in the case of auto MS/MS, the number of identified proteins hardly increases after the third run. It is cause to analyze the same precursor ion which is in high abundance repeatedly. Therefore, the effective analysis of low-abundance proteins such as those related to disease could be performed by repeating an LC/MS analysis with IBA (I).

The MS^3 analysis supplies the information of the peptide sequence, but lowers the total throughput. So, in this study, we set the conditions for running MS^3 analysis strictly; that

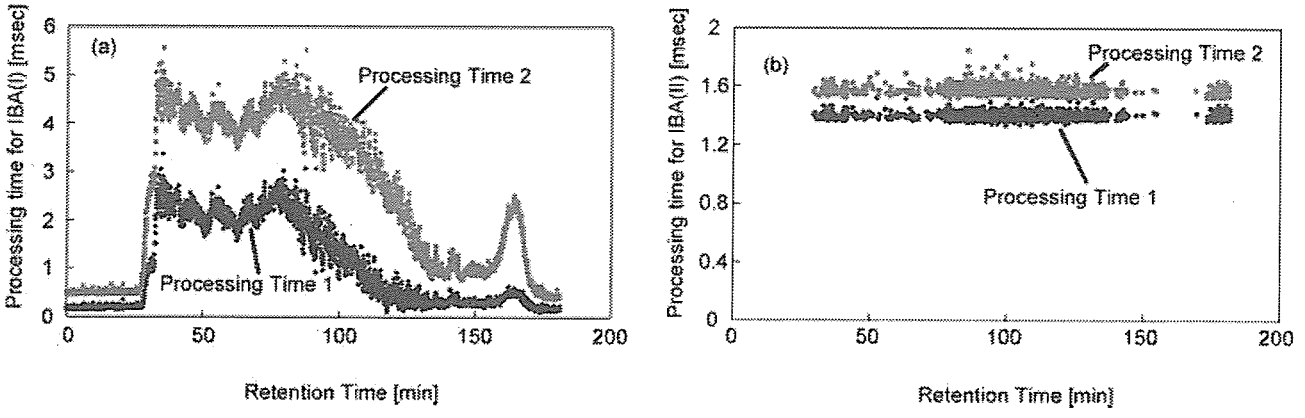


Figure 5. Processing time for real-time optimization. (a) IBA (I): Processing Time 1: time from noise filtering to peak detection, Processing Time 2: time from noise filtering, peak detection, inside database searching and selection of MS^2 precursor ion. (b) IBA (II): Processing Time 1: time from noise filtering to peak detection, Processing Time 2: time from noise filtering, peak detection, *de novo* sequencing and judgment of necessity of MS^3 analysis.

DOI: 10.1002/rcm

RCM

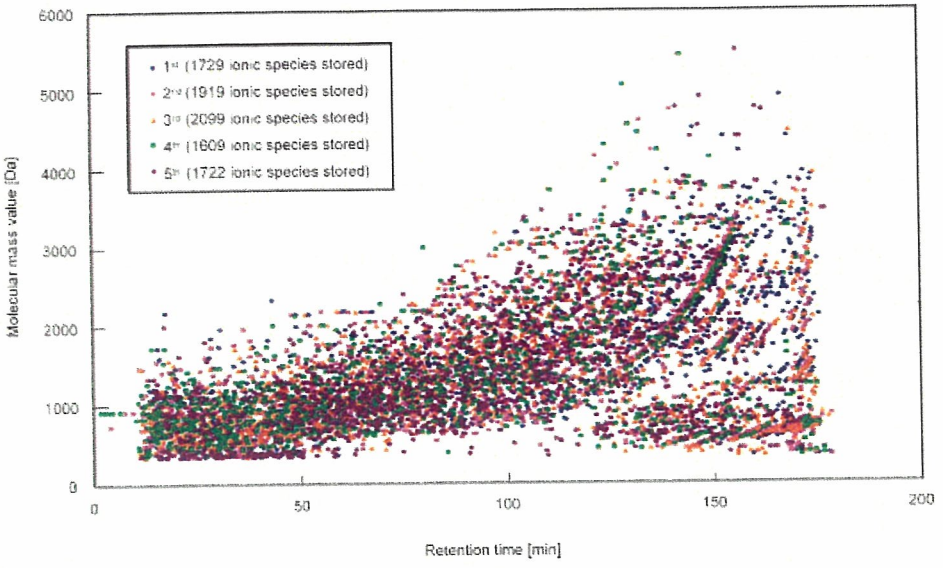


Figure 6. The precursor ions (molecular masses vs. retention times) selected by IBA (I) (5 cycles)

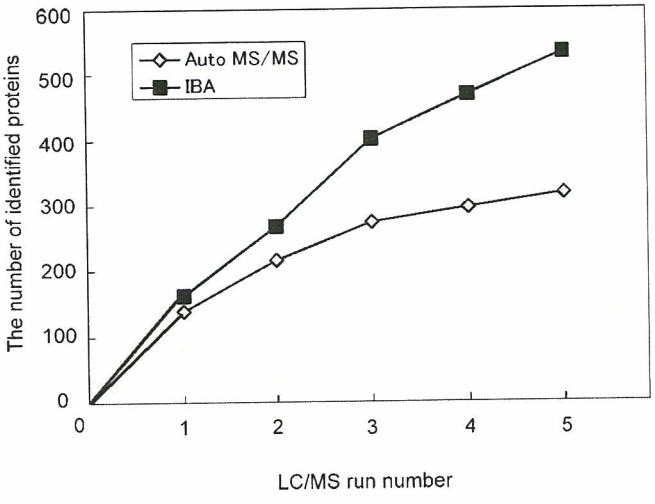


Figure 7. The number of identified proteins. Comparison of the number of identified proteins between IBA and auto MS/MS.

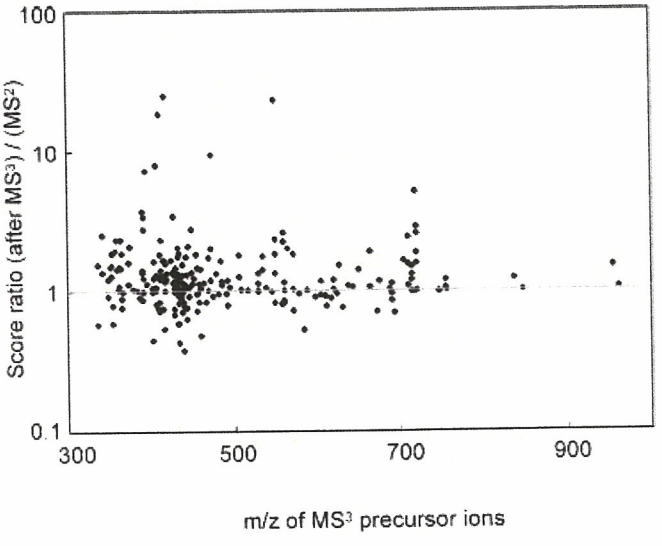


Figure 8. The score ratio (after MS^3/MS^2) of ion scores.

is, the threshold value of intensity of precursor ion to execute MS^3 analysis was set high. As a result, the frequency of MS^3 analysis depended on the LC/MS run and the ratio of the MS^3 analysis was from about 3 to 10% of the MS^2 analysis. Figure 8 shows the ratio of the Mascot® score of merged spectrum MS^2 with MS^3 to score of MS^2 spectrum. If the ratio is greater than 1.0, the identification accuracy has improved by the auto MS^3 analysis. In Figure 8, it was confirmed that peptide ion score was improved by the auto MS^3 analysis with 70% of the cases which executed MS^3. On the other hand, it was confirmed that the ratio was lower than 1.0 with 30% of the cases which performed MS^3. Figures 9 and 10 show typical examples of each case. In the case where the score of peptide improves, i.e. the ratio was greater than 1.0, it can be confirmed that the information on the amino acid sequence is supplemented by performing the MS^3 analysis to the y_{13}^{++} ion, as shown in Fig. 9. The improvement in the score is due to the increasing of the number of assigned fragment peaks with the auto MS^3 analysis.

On the other hand, Fig. 10 shows an example of the case where the score of the peptide decreases. In Fig. 10, the candidate peptide sequence at the 1st place has changed from DATKRR to DADGKLK with decreasing score. It is considered that the score has decreased because the candidate sequence DATKRR was a false-positive sequence. In the algorithm of Mascot®, the existence of unassigned peaks decreases the score. Therefore, in 30% of the cases which executed MS^3, it is considered that the score has decreased by the exclusion of the false-positive sequence and the identification accuracy of the protein improves by the exclusion of the false-positive candidates.

DOI: 10.1002/rcm