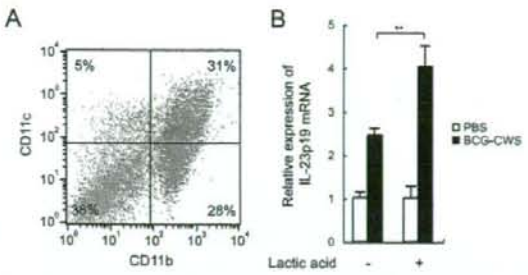

FIGURE 7. Lactic acid enhances the expression of *IL-23p19* mRNA in tumor-infiltrating immune cells. *A*, Tumor-infiltrating CD45-positive cells were purified from tumors formed by B16 mouse melanoma cells and stained for CD11b and CD11c. The stained cells were analyzed by FACS. *B*, One hundred thousand tumor-infiltrating immune cells (CD45-positive cells, 86%) were stimulated with 10 μg/ml BCG-CWS in the presence or absence of 15 mM lactic acid. After 4 h, the expression of *IL-23p19* and *β-actin* transcripts was measured by TaqMan RT-PCR. Relative expression was normalized to that of *β-actin*. The data represent mean values ± SD ($n = 3$). **, $p < 0.01$.

Lactic acid enhances the IL-23/IL-17 pathway

Because lactic acid further enhanced TLR-stimulated IL-23 production, we predicted that lactic acid should enhance IL-23 production from APCs stimulated with TLR ligand, leading to increased IL-17 production by IL-17-producing T cells. We then used OT-II-transgenic mice, which have $OVA_{323-339}$ peptide-specific T cells (25). In the presence of the OVA peptide, we examined the effect of lactic acid on the IL-17 production of OT-II mouse splenocytes that had been stimulated with the TLR ligand (PGN or BCG-CWS) for 4 days. Lactic acid remarkably enhanced the secretion of IL-17 induced by PGN and BCG-CWS (Fig. 6*A*, *left*, □ and ■). Anti-IL-23p19 Ab significantly inhibited the lactic acid-induced enhancement of IL-17 production (Fig. 6*A*, *left*, ■ and ▨). This result indicates that lactic acid indeed stimulates APCs to produce IL-23, which then drives peptide-activated T cells to produce IL-17. Interestingly, in the presence of the peptide, lactic acid induced IL-17 production even without the TLR ligand and this effect was inhibited by the anti-IL-23p19 Ab (Fig. 6, *A*, *left*, and *B*). The peptide or lactic acid alone rarely induced IL-17 production by splenocytes (Fig. 6*B*).

We also examined the effect of lactic acid on the secretion of IFN-γ (Fig. 6*A*, *right*). When the TLR ligand was present along with the peptide, the splenocytes produced IFN-γ. However, lactic acid did not potentiate IFN-γ production under these conditions (Fig. 6*A*, *right*, □ and ■). In contrast to IL-17 production, lactic acid did not induce IFN-γ production when the TLR ligand was absent (Fig. 6*A*, *right*). Notably, the effect of lactic acid on IFN-γ production was different when the TLR ligand was replaced by LPS. Although lactic acid potentiated IL-17 production by splenocytes irrespective of the ligand used for stimulation (Fig. 6*C*, *left*), it severely inhibited IFN-γ production (Fig. 6*C*, *right*).

Lactic acid enhances the expression of IL-23p19 *mRNA in tumor-infiltrating immune cells*

Many immune cells infiltrate the tumor microenvironment and induce local inflammation in and around the tumor. We examined whether tumor-infiltrating immune cells actually have the ability to express a high amount of IL-23 in response to lactic acid. We purified CD45-positive cells (CD11b- and/or CD11c-positive cells, 64%) that had infiltrated the tumors formed by B16 melanoma cells (Fig. 7*A*). These cells were stimulated with BCG-CWS and lactic acid for 4 h. The level of *IL-23p19* transcripts significantly increased in the presence of lactic acid (Fig. 7*B*). This result suggests that lactic acid also enhances the production of IL-23 in tumor-infiltrating immune cells, probably in macrophages and/or DCs.

Discussion

Tumor-secreted lactic acid enhances TLR ligand-dependent IL-23 expression in monocytes/macrophages

In this study, we sought to elucidate the mechanisms that induce immune cells in the tumor microenvironment to produce proinflammatory cytokines such as IL-23. We also asked, how does the tumor induce chronic inflammation without being attacked by the immune cells? We found that many tumor cell lines secrete factors that enhance TLR ligand-stimulated *IL-23p19* transcription and IL-23 secretion by human and mouse monocytes/macrophages (Fig. 1*A* and data not shown). We clarified that the main factor responsible for this effect is a molecule smaller than 500 Da that is protease resistant, heat stable, and only secreted from tumor cell lines when they are cultured in high glucose medium (Fig. 1, *D* and *E*, and data not shown). We then discovered that this molecule is lactic acid (Figs. 2–5). Lactic acid increased the production of proinflammatory cytokines, IL-23, and IL-6 from human monocytes (Fig. 5).

GM-CSF also enhanced TLR ligand-induced IL-23 and IL-12/23p40 production (Fig. 5). GM-CSF is secreted by many tumor cells as well as CADO-LC10 cells (35) and increases the expression of TLR2 (36). We suggest that tumor-secreted GM-CSF and lactic acid cooperate to increase IL-23 production following TLR2 ligand stimulation.

Predicted lactic acid signaling pathway

Although lactic acid is completely ionized, even under neutralized conditions (37), the neutralization of lactic acid suppressed its ability to enhance *IL-23p19* promoter activity in a pH-dependent manner (Fig. 4*A*). Moreover, neither sodium lactate nor acidification with hydrochloric acid enhanced *IL-23p19* transcription (Fig. 4, *B* and *D*). Lactate anions are transported together with protons through MCTs and a pH gradient is necessary for effective transport of lactic acid into cells (33). Therefore, we predict that lactic acid transported into monocytes/macrophages by MCTs may be recognized by an intracellular sensor that, along with the TLR signal, activates the *IL-23p19* promoter.

We found that DNA elements responsible for the enhancing activity of lactic acid are localized in the 2.7-kb 5′-flanking region of the human *IL-23p19* gene (Fig. 2*C*). We found four predicted NF-κB binding sites in this region (Fig. 1*C*). Although NF-κB signaling plays an important role in the TLR signaling pathway (4), lactic acid did not enhance the luciferase activity of a reporter vector carrying canonical NF-κB-binding elements (Fig. 2*D*). Therefore, the lactic acid signal pathway may be independent of the NF-κB pathway.

Lactic acid is not only a terminal metabolite of glycolysis but also a proinflammatory mediator

In most normal mammalian cells, the metabolism from glucose to lactate is inhibited by the presence of oxygen, which leads to the oxidization of pyruvate to CO_2 and H_2O in the mitochondria. However, in cancer cells, glycolysis is up-regulated, even in aerobic conditions, a phenomenon known as the "Warburg effect" or "aerobic glycolysis" (29, 30). This results in the production of large amounts of lactic acid and the accumulation of lactic acid in the microenvironment of many cancer cell types (31). High concentrations of lactate in solid tumors such as cervical carcinoma

(10.0 ± 2.9 μmol/g or >8.3 μmol/g) and head and neck cancer (>7.1 μmol/g) are associated with higher frequencies of distant metastasis (31, 38–40). Moreover, low lactate tumors in primary lesions are associated with longer disease-free survival than high lactate tumors. In addition, when the LDHA activity that induces the metabolism of pyruvate to lactate is suppressed, the tumorigenicity is severely diminished (41).

In this study, we showed that lactic acid enhances the expression of IL-23p19 in tumor-infiltrating immune cells activated by TLR stimuli (Fig. 7) and in splenocytes in the presence of Ag stimulus (data not shown) and induces the Ag- and IL-23-dependent secretion of IL-17 from splenocytes (Fig. 6). Significantly, we detected this IL-23-dependent enhancing activity even when TLR ligands were absent. Since lactic acid alone did not induce IL-23 production in the absence of Ag stimulation, we predicted that an interaction between APCs and T cells mediated by Ag and lactic acid causes IL-23-dependent IL-17 production in the absence of the TLR ligand. We observed that lactic acid severely inhibited IFN-γ production by Ag-stimulated splenocytes treated with LPS (Fig. 6*C*, *right*). The Kreutz group (42) reported that lactic acid represses the secretion of IL-12p70 by LPS-stimulated DCs. Therefore, lactic acid may inhibit IFN-γ production by suppressing the IL-12p70 production of monocytes/macrophages. In contrast, lactic acid did not affect IFN-γ production of Ag-stimulated splenocytes treated with PGN or BCG-CWS (Fig. 6*A*, *right*), because it is assumed that PGN and BCG-CWS do not stimulate IL-12p70 production by monocytes/macrophages (14). These results suggest that upon antigenic stimulation of T cells, lactic acid mediates the activation of the IL-23/IL-17 pathway rather than the induction of IFN-γ-producing Th1 cells. We predict that lactic acid derived from tumor cells may induce inflammation but not the infiltration of CTLs in the tumor microenvironment, even in the absence of TLR ligand stimuli from microbial infections, and that inflammation promotes angiogenesis and tumor development.

The Kreutz group (42, 43) also reported that lactic acid downregulates the cytokine production and proliferation of CTLs. Furthermore, the Murray group (44) showed that MCT1 inhibitors, which inhibit the transport of lactic acid, suppress T lymphocyte proliferation. These groups suggested that a high concentration of extracellular lactic acid or inhibition of lactic acid excretion might cause intracellular accumulation of lactic acid with consequent disturbance of T cell metabolism and function. In contrast, our results indicate that lactic acid up-regulates *IL-23p19* transcription in monocytes/macrophages.

In conclusion, our results show that lactic acid acts as a novel tumor-derived factor that strongly induces the IL-23/IL-17 proinflammatory pathway without inducing Th1 responses. Thus, the production and excretion of lactic acid appears to be not only essential for the proliferation of tumor cells with up-regulated glycolysis, but also for the induction of inflammation in the tumor microenvironment, which promotes tumor progression. Therefore, lactic acid and the lactic acid/IL-23 signal pathway may be an attractive target for treating tumors.

Acknowledgments

We thank K. Kawata, T. Yasuda, and C. Kozai in our laboratory for technical assistance and Dr. M. Saio (Gifu University, Gifu, Japan) for technical support. We also thank Drs. K. Toyoshima (RIKEN, Yokohama, Japan) and I. Azuma (Muroran Institute of Technology, Muroran, Japan) for providing BCG-CWS, Dr. W. R. Heath (The Walter and Eliza Hall Institute of Medical Research, Parkville, Australia) for providing OT-II mice, and Drs. J. Takeda and K. Yusa (Osaka University, Suita, Japan) for providing pGKEM7-neoW.

Disclosures

The authors have no financial conflict of interest.

References

1. de Visser, K. E., A. Eichten, and L. M. Coussens. 2006. Paradoxical roles of the immune system during cancer development. *Nat. Rev. Cancer* 6: 24–37.
2. Karin, M., T. Lawrence, and V. Nizet. 2006. Innate immunity gone awry: linking microbial infections to chronic inflammation and cancer. *Cell* 124: 823–835.
3. Balkwill, F., K. A. Charles, and A. Mantovani. 2005. Smoldering and polarized inflammation in the initiation and promotion of malignant disease. *Cancer Cell* 7: 211–217.
4. Akira, S., S. Uematsu, and O. Takeuchi. 2006. Pathogen recognition and innate immunity. *Cell* 124: 783–801.
5. Oppmann, B., R. Lesley, B. Blom, J. C. Timans, Y. Xu, B. Hunte, F. Vega, N. Yu, J. Wang, K. Singh, et al. 2000. Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. *Immunity* 13: 715–725.
6. Hunter, C. A. 2005. New IL-12-family members: IL-23 and IL-27, cytokines with divergent functions. *Nat. Rev. Immunol.* 5: 521–531.
7. Veldhoen, M., R. J. Hocking, C. J. Atkins, R. M. Locksley, and B. Stockinger. 2006. TGF-β in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity* 24: 179–189.
8. Bettelli, E., Y. Carrier, W. Gao, T. Korn, T. B. Strom, M. Oukka, H. L. Weiner, and V. K. Kuchroo. 2006. Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* 441: 235–238.
9. Mangan, P. R., L. E. Harrington, D. B. O'Quinn, W. S. Helms, D. C. Bullard, C. O. Elson, R. D. Hatton, S. M. Wahl, T. R. Schoeb, and C. T. Weaver. 2006. Transforming growth factor-β induces development of the T_H17 lineage. *Nature* 441: 231–234.
10. Weaver, C. T., L. E. Harrington, P. R. Mangan, M. Gavrieli, and K. M. Murphy. 2006. Th17: an effector CD4 T cell lineage with regulatory T cell ties. *Immunity* 24: 677–688.
11. Aggarwal, S., N. Ghilardi, M. H. Xie, F. J. de Sauvage, and A. L. Gurney. 2003. Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17. *J. Biol. Chem.* 278: 1910–1914.
12. Zheng, Y., D. M. Danilenko, P. Valdez, I. Kasman, J. Eastham-Anderson, J. Wu, and W. Ouyang. 2007. Interleukin-22, a T_H17 cytokine, mediates IL-23-induced dermal inflammation and acanthosis. *Nature* 445: 648–651.
13. Begum, N. A., K. Ishii, M. Kurita-Taniguchi, M. Tanabe, M. Kobayashi, Y. Moriwaki, M. Matsumoto, Y. Fukumori, I. Azuma, K. Toyoshima, and T. Seya. 2004. *Mycobacterium bovis* BCG cell wall-specific differentially expressed genes identified by differential display and cDNA subtraction in human macrophages. *Infect. Immun.* 72: 937–948.
14. McKenzie, B. S., R. A. Kastelein, and D. J. Cua. 2006. Understanding the IL-23-IL-17 immune pathway. *Trends Immunol.* 27: 17–23.
15. Tsuji, S., M. Matsumoto, O. Takeuchi, S. Akira, I. Azuma, A. Hayashi, K. Toyoshima, and T. Seya. 2000. Maturation of human dendritic cells by cell wall skeleton of *Mycobacterium bovis* bacillus Calmette-Guérin: involvement of toll-like receptors. *Infect. Immun.* 68: 6883–6890.
16. Langowski, J. L., X. Zhang, L. Wu, J. D. Mattson, T. Chen, K. Smith, B. Basham, T. McClanahan, R. A. Kastelein, and M. Oft. 2006. IL-23 promotes tumour incidence and growth. *Nature* 442: 461–465.
17. Kryczek, I., S. Wei, L. Zou, S. Altuwaijri, W. Szeliga, J. Kolls, A. Chang, and W. Zou. 2007. Cutting edge: Th17 and regulatory T cell dynamics and the regulation by IL-2 in the tumor microenvironment. *J. Immunol.* 178: 6730–6733.
18. Kolls, J. K., and A. Linden. 2004. Interleukin-17 family members and inflammation. *Immunity* 21: 467–476.
19. Takahashi, H., M. Numasaki, M. T. Lotze, and H. Sasaki. 2005. Interleukin-17 enhances bFGF-, HGF- and VEGF-induced growth of vascular endothelial cells. *Immunol. Lett.* 98: 189–193.
20. Numasaki, M., J. Fukushi, M. Ono, S. K. Narula, P. J. Zavodny, T. Kudo, P. D. Robbins, H. Tahara, and M. T. Lotze. 2003. Interleukin-17 promotes angiogenesis and tumor growth. *Blood* 101: 2620–2627.
21. Matsumoto, M., T. Seya, S. Kikkawa, S. Tsuji, K. Shida, M. Nomura, M. Kurita-Taniguchi, H. Ohigashi, H. Yokouchi, K. Takami, et al. 2001. Interferon γ-producing ability in blood lymphocytes of patients with lung cancer through activation of the innate immune system by BCG cell wall skeleton. *Int. Immunopharmacol.* 1: 1559–1569.
22. Sakuma, T., K. Kodama, T. Hara, Y. Eshita, N. Shibata, M. Matsumoto, T. Seya, and Y. Mori. 1993. Levels of complement regulatory molecules in lung cancer: disappearance of the D17 epitope of CD55 in small-cell carcinoma. *Jpn. J. Cancer Res.* 84: 753–759.
23. Akazawa, T., H. Masuda, Y. Saeki, M. Matsumoto, K. Takeda, K. Tsujimura, K. Kuzushima, T. Takahashi, I. Azuma, S. Akira, et al. 2004. Adjuvant-mediated tumor regression and tumor-specific cytotoxic response are impaired in MyD88-deficient mice. *Cancer Res.* 64: 757–764.
24. Azuma, I., E. E. Ribi, T. J. Meyer, and B. Zbar. 1974. Biologically active components from mycobacterial cell walls: I. Isolation and composition of cell wall skeleton and component P3. *J. Natl. Cancer Inst.* 52: 95–101.
25. Barnden, M. J., J. Allison, W. R. Heath, and F. R. Carbone. 1998. Defective TCR expression in transgenic mice constructed using cDNA-based α- and β-chain genes under the control of heterologous regulatory elements. *Immunol. Cell Biol.* 76: 34–40.
26. Livak, K. J., and T. D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta CT}$ method. *Methods* 25: 402–408.

27. Wilson, M. C., V. N. Jackson, C. Heddle, N. T. Price, H. Pilegaard, C. Juel, A. Bonen, I. Montgomery, O. F. Hutter, and A. P. Halestrap. 1998. Lactic acid efflux from white skeletal muscle is catalyzed by the monocarboxylate transporter isoform MCT3. *J. Biol. Chem.* 273: 15920–15926.
28. Saio, M., S. Radoja, M. Marino, and A. B. Frey. 2001. Tumor-infiltrating macrophages induce apoptosis in activated $CD8^+$ T cells by a mechanism requiring cell contact and mediated by both the cell-associated form of TNF and nitric oxide. *J. Immunol.* 167: 5583–5593.
29. Kim, J. W., and C. V. Dang. 2006. Cancer's molecular sweet tooth and the Warburg effect. *Cancer Res.* 66: 8927–8930.
30. Gatenby, R. A., and R. J. Gillies. 2004. Why do cancers have high aerobic glycolysis? *Nat. Rev. Cancer* 4: 891–899.
31. Walenta, S., and W. F. Mueller-Klieser. 2004. Lactate: mirror and motor of tumor malignancy. *Semin. Radiat. Oncol.* 14: 267–274.
32. Heinemeyer, T., E. Wingender, I. Reuter, H. Hermjakob, A. E. Kel, O. V. Kel, E. V. Ignatieva, E. A. Ananko, O. A. Podkolodnaya, F. A. Kolpakov, et al. 1998. Databases on transcriptional regulation: TRANSFAC, TRRD, and COMPEL. *Nucleic Acids Res.* 26: 362–367.
33. Halestrap, A. P., and N. T. Price. 1999. The proton-linked monocarboxylate transporter (MCT) family: structure, function, and regulation. *Biochem. J.* 343: 281–299.
34. Trinchieri, G. 2003. Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nat. Rev. Immunol.* 3: 133–146.
35. Bronte, V., D. B. Chappell, E. Apolloni, A. Cabrelle, M. Wang, P. Hwu, and N. P. Restifo. 1999. Unopposed production of granulocyte-macrophage colony-stimulating factor by tumors inhibits $CD8^+$ T cell responses by dysregulating antigen-presenting cell maturation. *J. Immunol.* 162: 5728–5737.
36. Flo, T. H., O. Halaas, S. Torp, L. Ryan, E. Lien, B. Dybdahl, A. Sundan, and T. Espevik. 2001. Differential expression of Toll-like receptor 2 in human cells. *J. Leukocyte Biol.* 69: 474–481.
37. Gladden, L. B. 2004. Lactate metabolism: a new paradigm for the third millennium. *J. Physiol.* 558: 5–30.
38. Walenta, S., M. Wetterling, M. Lehrke, G. Schwickert, K. Sundfor, E. K. Rofstad, and W. Mueller-Klieser. 2000. High lactate levels predict likelihood of metastases, tumor recurrence, and restricted patient survival in human cervical cancers. *Cancer Res.* 60: 916–921.
39. Schwickert, G., S. Walenta, K. Sundfor, E. K. Rofstad, and W. Mueller-Klieser. 1995. Correlation of high lactate levels in human cervical cancer with incidence of metastasis. *Cancer Res.* 55: 4757–4759.
40. Walenta, S., A. Salameh, H. Lyng, J. F. Evensen, M. Mitze, E. K. Rofstad, and W. Mueller-Klieser. 1997. Correlation of high lactate levels in head and neck tumors with incidence of metastasis. *Am. J. Pathol.* 150: 409–415.
41. Fantin, V. R., J. St.-Pierre, and P. Leder. 2006. Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. *Cancer Cell* 9: 425–434.
42. Gottfried, E., L. A. Kunz-Schughart, S. Ebner, W. Mueller-Klieser, S. Hoves, R. Andreesen, A. Mackensen, and M. Kreutz. 2006. Tumor-derived lactic acid modulates dendritic cell activation and antigen expression. *Blood* 107: 2013–2021.
43. Fischer, K., P. Hoffmann, S. Voelkl, N. Meidenbauer, J. Ammer, M. Edinger, E. Gottfried, S. Schwarz, G. Rothe, S. Hoves, et al. 2007. Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood* 109: 3812–3819.
44. Murray, C. M., R. Hutchinson, J. R. Bantick, G. P. Belfield, A. D. Benjamin, D. Brazma, R. V. Bundick, I. D. Cook, R. I. Craggs, S. Edwards, et al. 2005. Monocarboxylate transporter MCT1 is a target for immunosuppression. *Nat. Chem. Biol.* 1: 371–376.

JOURNAL OF BACTERIOLOGY, May 2008, p. 3613–3621 Vol. 190, No. 10
0021-9193/08/$08.00+0 doi:10.1128/JB.01850-07

Structural Analysis and Biosynthesis Gene Cluster of an Antigenic Glycopeptidolipid from *Mycobacterium intracellulare*▽†

Nagatoshi Fujiwara,[1]* Noboru Nakata,[2] Takashi Naka,[1,3] Ikuya Yano,[3] Matsumi Doe,[4] Delphi Chatterjee,[5] Michael McNeil,[5] Patrick J. Brennan,[5] Kazuo Kobayashi,[6] Masahiko Makino,[2] Sohkichi Matsumoto,[1] Hisashi Ogura,[7] and Shinji Maeda[8]

Department of Host Defense[1] and Virology,[7] Osaka City University Graduate School of Medicine, Osaka 545-8585, Japan; Department of Microbiology, Leprosy Research Center, National Institute of Infectious Diseases, Tokyo 189-0002, Japan[2]; Japan BCG Laboratory, Tokyo 204-0022, Japan[3]; Department of Chemistry, Graduate School of Science, Osaka City University, Osaka 558-8585, Japan[4]; Department of Microbiology, Immunology and Pathology, Colorado State University, Colorado 80523[5]; Department of Immunology, National Institute of Infectious Diseases, Tokyo 162-8640, Japan[6]; and Molecular Epidemiology Division, Mycobacterium Reference Center, The Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Tokyo 204-8533, Japan[8]

Received 24 November 2007/Accepted 1 March 2008

***Mycobacterium avium-Mycobacterium intracellulare* complex (MAC) is the most common isolate of nontuberculous mycobacteria and causes pulmonary and extrapulmonary diseases. MAC species can be grouped into 31 serotypes by the epitopic oligosaccharide structure of the species-specific glycopeptidolipid (GPL) antigen. The GPL consists of a serotype-common fatty acyl peptide core with 3,4-di-*O*-methyl-rhamnose at the terminal alaninol and a 6-deoxy-talose at the *allo*-threonine and serotype-specific oligosaccharides extending from the 6-deoxy-talose. Although the complete structures of 15 serotype-specific GPLs have been defined, the serotype 16-specific GPL structure has not yet been elucidated. In this study, the chemical structure of the serotype 16 GPL derived from *M. intracellulare* was determined by using chromatography, mass spectrometry, and nuclear magnetic resonance analyses. The result indicates that the terminal carbohydrate epitope of the oligosaccharide is a novel *N*-acyl-dideoxy-hexose. By the combined linkage analysis, the oligosaccharide structure of serotype 16 GPL was determined to be 3-2′-methyl-3′-hydroxy-4′-methoxy-pentanoyl-amido-3,6-dideoxy-β-hexose-(1→3)-4-*O*-methyl-α-L-rhamnose-(1→3)-α-L-rhamnose-(1→3)-α-L-rhamnose-(1→2)-6-deoxy-α-L-talose. Next, the 22.9-kb serotype 16-specific gene cluster involved in the glycosylation of oligosaccharide was isolated and sequenced. The cluster contained 17 open reading frames (ORFs). Based on the similarity of the deduced amino acid sequences, it was assumed that the ORF functions include encoding three glycosyltransferases, an acyltransferase, an aminotransferase, and a methyltransferase. An *M. avium* serotype 1 strain was transformed with cosmid clone no. 253 containing *gtfB-drrC* of *M. intracellulare* serotype 16, and the transformant produced serotype 16 GPL. Together, the ORFs of this serotype 16-specific gene cluster are responsible for the biosynthesis of serotype 16 GPL.**

Mycobacterial diseases, such as tuberculosis and infection due to nontuberculous mycobacteria (NTM), are still among the most serious infectious diseases in the world. The incidence is increasing because of the spread of drug-resistant mycobacteria and the human immunodeficiency virus (HIV) infection/AIDS epidemic (16, 17, 30). *Mycobacterium avium-Mycobacterium intracellulare* complex (MAC) is the most common among isolates of NTM and is distributed ubiquitously in the environment. MAC causes pulmonary and extrapulmonary diseases in both immunocompromised and immunocompetent hosts. It affects primarily patients with advanced HIV infection. MAC includes at least two mycobacterial species, *M. avium* and *M. intracellulare*, that cannot be differentiated on the basis of traditional physical and biochemical tests (1, 41).

The cell envelope of mycobacteria is a complex and unusual structure. The key feature of this structure is an extraordinarily high lipid concentration (6, 10). To better understand the pathogenesis of MAC infection, it is necessary to elucidate the molecular structure and biochemical features of the lipid components. Among MAC lipids, the glycopeptidolipid (GPL) is of particular importance, because it shows not only serotype-specific antigenicity but also immunomodulatory activities in the host immune responses (2, 9, 23). Structurally, GPLs are composed of two parts, a tetrapeptide-amino alcohol core and a variable oligosaccharide (OSE). C_{26}-C_{34} fatty acyl-D-phenylalanine-D-*allo*-threonine-D-alanine-L-alaninol (D-Phe-D-*allo*-Thr-D-Ala-L-alaninol) is further linked with 6-deoxy talose (6-d-Tal) and 3,4-di-*O*-methyl rhamnose (3,4-di-*O*-Me-Rha) at D-*allo*-Thr and the terminal L-alaninol, respectively. This type of core GPL is found in all subspecies of MAC, shows a common antigenicity, and is further glycosylated at 6-d-Tal to form a serotype-specific OSE.

At present, 31 distinct serotype-specific GPLs have been identified serologically and chromatographically (9). Although the standard technique for differentiation of MAC subspecies

* Corresponding author. Mailing address: Department of Host Defense, Osaka City University Graduate School of Medicine, 1-4-3 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan. Phone: 81 6 6645 3746. Fax: 81 6 6645 3747. E-mail: fujiwara@med.osaka-cu.ac.jp.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 7 March 2008.

has been serotyping based on the OSE residue of its GPL, the complete structures of only 15 GPLs have been defined. In addition to the chemical structures of various GPLs, genes encoding the glycosylation pathways in the biosynthesis of GPL have been identified and characterized (12, 21, 31). Epidemiological studies have shown that MAC serotypes 4 and 8 are the most frequently isolated from patients, and MAC serotype 16 is one of the next most common groups (32, 40). It has been suggested that the serotypes of MAC isolates participate in their virulence (29), and thus, understanding of the structure-pathogenicity relationship of GPLs is necessary. In the present study, we demonstrate the complete OSE structure of the GPL derived from serotype 16 MAC (*M. intracellulare*), which has a unique terminal-acylated-amido sugar, and we characterized the serotype 16 GPL-specific gene cluster involved in the glycosylation of carbohydrates.

MATERIALS AND METHODS

Bacterial strains and preparation of GPL. *M. intracellulare* serotype 16 strain ATCC 13950^T (NF 115) was purchased from the American Type Culture Collection (Manassas, VA). Three clinical isolates of *M. intracellulare* serotype 16 (NF 116 and 117) and *M. avium* serotype 1 (NF 113) were maintained in The Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association. The preparation of GPL was performed as described previously (18, 24, 26). Briefly, each strain of *M. intracellulare* serotype 16 was grown in Middlebrook 7H9 broth (Difco Laboratories, Detroit, MI) with 0.5% glycerol and 10% Middlebrook oleic acid-albumin-dextrose-catalase enrichment (Difco) at 37°C for 2 to 3 weeks. The heat-killed bacteria were sonicated, and crude lipids were extracted with chloroform-methanol (2:1, vol/vol). The extracted lipids were dried and hydrolyzed with 0.2 N sodium hydroxide in methanol at 37°C for 2 h. After neutralization with 6 N hydrochloric acid, alkaline-stable lipids were partitioned by a two-layer system of chloroform-methanol (2:1, vol/vol) and water. The organic phase was recovered, evaporated, and precipitated with acetone to remove any acetone-insoluble components containing phospholipids and glycolipids. The supernatant was collected by centrifugation, dried, and then treated with a Sep-Pak silica cartridge (Waters Corporation, Milford, MA) with washing (chloroform-methanol, 95:5, vol/vol) and elution (chloroform-methanol, 1:1, vol/vol) for partial purification. GPL was completely purified by preparative thin-layer chromatography (TLC) of Silica Gel G (20 by 20 cm, 250 μm; Uniplate; Analtech, Inc., Newark, DE). The TLC plate was repeatedly developed with chloroform-methanol-water (65:25:4 and 60:16:2, vol/vol/vol) until a single spot was obtained. After exposure of the TLC plate to iodine vapor, the GPL band was marked, and then, the silica gels were scraped off and the GPL was eluted with chloroform-methanol (2:1, vol/vol).

Preparation of OSE moiety. β elimination of GPL was performed with alkaline borohydride, and the OSE elongated from D-*allo*-Thr was released as described previously (18, 24). Briefly, the GPL was dissolved in ethanol, and an equal volume of 10 mg/ml sodium borohydride or borodeuteride in 0.5 N sodium hydroxide was added and then stirred at 60°C for 16 h. The reaction mixture was decationized with Dowex 50W-X8 beads (Dow Chemical Company, Midland, MI), collected, and evaporated under nitrogen to remove boric acid. The dried residue was partitioned in two layers of chloroform-methanol (2:1, vol/vol) and water. The upper aqueous phase was recovered and evaporated. In these processes, the serotype 16-specific OSE was purified as an oligoglycosyl alditol.

MALDI-TOF and MALDI-TOF/TOF MS analyses. The molecular species of the intact GPL was detected by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) with an Ultraflex II (Bruker Daltonics, Billerica, MA). The GPL was dissolved in chloroform-methanol (2:1, vol/vol) at a concentration of 1 mg/ml, and 1 μl was applied directly to the sample plate, and then 1 μl of 10 mg/ml 2,5-dihydroxybenzoic acid in chloroform-methanol (1:1, vol/vol) was added as a matrix. The intact GPL was analyzed in the reflectron mode with an accelerating voltage operating in a positive mode of 20 kV (5). Then the fragment pattern of the OSE was analyzed with MALDI-TOF/TOF MS. The OSE was dissolved in ethanol-water (3:7, vol/vol), and the matrix was 10 mg/ml 2,5-dihydroxybenzoic acid in ethanol-water (3:7, vol/vol). The OSE and the matrix were applied to the sample plate according to the method for intact GPL and analyzed in the lift-lift mode.

GC and GC-MS analyses of carbohydrates and *N*-acylated short-chain fatty acid. To determine the glycosyl composition and linkage position, gas chromatography (GC) and GC-MS analyses of partially methylated alditol acetate derivatives were performed. Perdeuteromethylation was conducted by the modified procedure of Hakomori as described previously (18, 20). Briefly, the dried OSE was dissolved with a mixture of dimethyl sulfoxide and sodium hydroxide, and deuteromethyl iodide was added. The reaction mixture was stirred at room temperature for 15 min and then water and chloroform were added. The chloroform-containing perdeuteromethylated OSE layer was collected, washed with water two times, and then completely evaporated. Partially deuteromethylated alditol acetates were prepared from perdeuteromethylated OSE by hydrolysis with 2 N trifluoroacetic acid at 120°C for 2 h, reduction with 10 mg/ml sodium borodeuteride at 25°C for 2 h, and acetylation with acetic anhydride at 100°C for 1 h (8, 18, 25). To identify amino-linked fatty acids, acidic methanolysis of serotype 16 GPL was performed with 1.25 M hydrogen chloride in methanol (Sigma-Aldrich, St. Louis, MO) at 100°C for 90 min, and the fatty acid methyl esters were extracted with *n*-hexane under the cooled ice. GC was performed using a 5890 series II gas chromatograph (Hewlett Packard, Avondale, PA) equipped with a fused SPB-1 capillary column (30 m, 0.25-mm inner diameter; Supelco Inc., Bellefonte, PA). Helium was used for electron impact (EI)-MS and isobutane for chemical ionization (CI)-MS as a carrier gas. A JMS SX102A double-focusing mass spectrometer (JEOL, Tokyo, Japan) was connected to the gas chromatograph as a mass detector. The molecular separator and the ion source energy were 70 eV for EI and 30 eV for CI, and the accelerating voltage was 8 kV. The D and L configurations of Rha residues were determined by comparative GC-MS analysis of trimethylsilylated (*S*)-(+)-sec-butyl glycosides and (*R*)-(−)-sec-butyl glycosides prepared from an authentic standard L-Rha (19).

NMR analysis of GPL. The GPL was dissolved in chloroform-d ($CDCl_3$)-methanol-d_4 (CD_3OD) (2:1, vol/vol). To define the anomeric configurations of each glycosyl residue, ^1H and ^{13}C nuclear magnetic resonance (NMR) was employed. Both homonuclear correlation spectrometry (COSY) and ^1H-detected [^1H, ^{13}C] heteronuclear multiple-quantum correlation (HMQC) were recorded with a Bruker Avance-600 (Bruker BioSpin Corp., Billerica, MA), as described previously (9, 18, 24, 34).

Construction of *M. intracellulare* serotype 16 cosmid library. A cosmid library of *M. intracellulare* serotype 16 strain ATCC 13950^T was constructed as described previously (18). Bacterial cells were disrupted mechanically, and genomic DNA was extracted with phenol-chloroform and then precipitated with ethanol. Genomic DNA randomly sheared into 30- to 50-kb fragments in the extraction process was fractionated and electroeluted from agarose gels using a Takara Recochip (Takara, Kyoto, Japan). These DNA fragments were rendered blunt ended using T4 DNA polymerase and deoxynucleoside triphosphates and then were ligated to dephosphorylated arms of pYUB412 (XbaI-EcoRV and EcoRV-XbaI), which were the kind gifts of William R. Jacobs, Jr. (Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY). The cosmid vector pYUB412 is an *Escherichia coli-Mycobacterium* shuttle vector with the *int-attP* sequence for integration into a mycobacterial chromosome, *oriE* for replication in *E. coli*, a hygromycin resistance gene, and an ampicillin resistance gene. After in vitro packaging using Gigapack III Gold extracts (Stratagene, La Jolla, CA), recombinant cosmids were introduced into *E. coli* STBL2 [F$^-$ *mcrA* Δ(*mcrBC-hsdRMS-mrr*) *recA1 endA1 lon gyrA96 thi supE44 relA1* λ Δ(*lac-proAB*)] and stored at −80°C in 50% glycerol.

Isolation of cosmid clones carrying biosynthesis gene cluster of serotype 16 GPL and sequence analysis. Isolation of DNA from *E. coli* transductants was performed as described by Supply et al., with modifications (39). The colonies were picked, transferred to a 1.5-ml tube containing 50 μl of water, and then heated at 98°C for 5 min. After centrifugation at 14,000 rpm for 5 min, the supernatant was used as the PCR template. PCR was used to isolate cosmid clones carrying the rhamnosyltransferase (*rtfA*) gene with primers rtfA-F (5′-TTTGGAGCGACGAGTTCATC-3′) and rtfA-R (5′-GTGTAGTTGACCACG CCGAC-3′). *rtfA* encodes an enzyme responsible for the transfer of Rha to 6-d-Tal in OSE (14, 31). The insert of cosmid clone no. 253 was sequenced using a BigDye Terminator, version 3.1, cycle sequencing kit (Applied Biosystems, Foster City, CA) and an ABI Prism 310 gene analyzer (Applied Biosystems). The putative function of each open reading frame (ORF) was identified by similarity searches between the deduced amino acid sequences and known proteins using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and FramePlot (http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl) with the DNASIS computer program (Hitachi Software Engineering, Yokohama, Japan).

Transformation of *M. avium* serotype 1 strain with cosmid clone no. 253. An *M. avium* serotype 1 strain (NF113) was transformed with pYUB412-cosmid clone no. 253 by electroporation, and hygromycin-resistant colonies were iso-

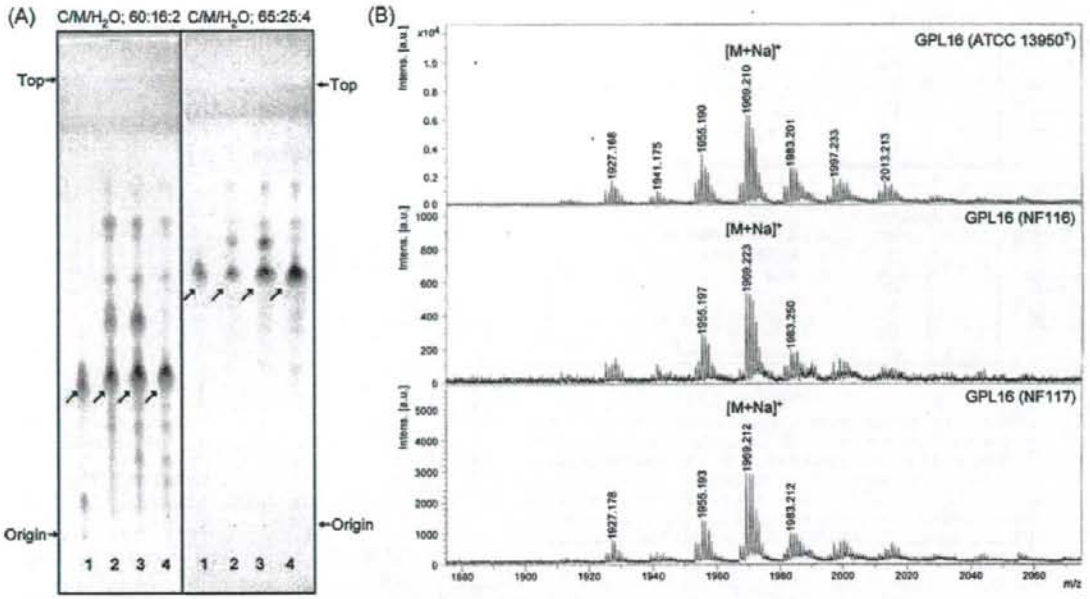

FIG. 1. TLC patterns and MALDI-TOF MS spectra of serotype 16 GPL. (A) Serotype 16 GPL purified from *M. intracellulare* ATCC 13950^T (NF 115) and the alkaline-stable lipids derived from ATCC 13950^T and two clinical isolates (NF 116 and 117) from left to right were developed on TLC plates with solvent systems of chloroform-methanol-water (65:25:4 and 60:16:2, vol/vol/vol). (B) The MALDI-TOF MS spectra were acquired using 10 mg/ml 2,5-dihydroxybenzoic acid in chloroform-methanol (1:1, vol/vol) as a matrix, and the molecularly related ions were detected as $[M+Na]^+$ in positive mode. Intens., intensity; a.u., absorbance units.

lated. Alkaline-stable lipids were prepared, and productive GPLs were examined by TLC and MALDI-TOF MS analyses.

Nucleotide sequence accession number. The nucleotide sequence reported here has been deposited in the NCBI GenBank database under accession no. AB355138.

RESULTS

Purification and molecular weight of intact GPL. Serotype 16 GPL from *M. intracellulare* ATCC 13950^T (NF 115) was detected as a spot by TLC, and the R_f values were 0.35 and 0.56 when developed with chloroform-methanol-water (60:16:2 and 65:25:4, vol/vol/vol, respectively). Two clinical isolates of *M. intracellulare*, NF 116 and 117, had serotype 16 GPLs that showed the same R_f values as the serotype 16 GPL derived from strain ATCC 13950^T. The serotype 16 GPL of *M. intracellulare* strain ATCC 13950^T was purified repeatedly by TLC and was shown as a single spot by TLC (Fig. 1A). The MALDI-TOF MS spectra of each serotype 16 GPL showed *m/z* 1969 for $[M+Na]^+$ as the main molecularly related ion in positive mode, with the homologous ions differing by 14 mass units at 1,955 and 1,983 (Fig. 1B). As a result, the main molecular weight of serotype 16 GPL was 1,946, which implied that it has a novel carbohydrate chain elongated from D-*allo*-Thr.

Carbohydrate composition of serotype 16 OSE. To determine the glycosyl compositions of serotype 16 OSE, alditol acetate derivatives of the serotype 16 GPL were analyzed by GC and GC-MS. The structurally defined serotype 4 GPL was used as a reference standard (9, 35). Comparison of the retention time and GC mass spectra (Fig. 2) with the alditol acetate derivatives of the serotype 16 GPL showed the presence of 3,4-di-*O*-Me-Rha, 4-*O*-Me-Rha, Rha, 6-d-Tal, and an unknown sugar residue (X1) in a ratio of approximately 1:1:2:1:1. The alditol acetate of X1 was eluted at a retention time of 29.3 min, greater than that of glucitol acetate on the SPB-1 column. The CI-MS spectrum of X1 was $[M+H]^+$ at *m/z* 520 as a parent ion and *m/z* 460 as a loss of 60 (acetate). The fragment ions of X1 sugar showed characteristic patterns in EI-MS. *m/z* 360 indicated the cleavage of C-3 and C-4, and *m/z* 300, 240, and 180 were fragmented with a loss of 60 (acetate). Similarly, *m/z* 374 indicated the cleavage of C-2 and C-3, and *m/z* 314 and 254 were fragmented with a loss of 60 (Fig. 3A and B). These results indicated that X1 was 3,6-dideoxy hexose (Hex). The odd molecular weight of X1, 519, and *m/z* 187, 127, and 59 implied the presence of one amido group esterified with a short-chain fatty acid, possibly. After methanolysis of serotype 16 GPL, the resultant fatty acid methyl esters were extracted carefully and analyzed by GC-MS. The EI-MS spectrum of a short-chain fatty acid methyl ester showed mass ions at *m/z* 176 ($[M]^+$), 145 ($[M-31]^+$), 117 ($[M-59]^+$), 99, 88, 85, and 59 (Fig. 3C) (33, 37). Taking the results together, X1 was structurally determined to be 3-2′-methyl-3′-hydroxy-4′-methoxy-pentanoyl-amido-3,6-dideoxy-Hex.

Glycosyl linkage and sequence of serotype 16 OSE. To determine the glycosyl linkage and sequence of the OSE, GC-MS of perdeuteromethylated alditol acetates and MALDI-TOF/TOF MS of the oligoglycosyl alditol from serotype 16 OSE

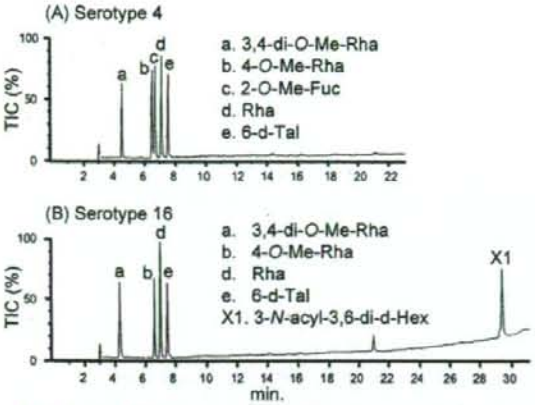

FIG. 2. Gas chromatograms of the alditol acetate derivatives from serotype 4 (A) and serotype 16 (B) GPLs. Total ion chromatograms (TIC) are shown. GC was performed on an SPB-1-fused silica column with a temperature program of 160°C for 2 min, followed by an increase of 4°C/min to 220°C, and holding at 220°C for 15 min. Comparison to the GC spectrum of serotype 4 GPL shows that serotype 16 GPL is composed of 3,4-di-*O*-Me-Rha, 4-*O*-Me-Rha, Rha, 6-d-Tal, and an unknown X1 sugar residue.

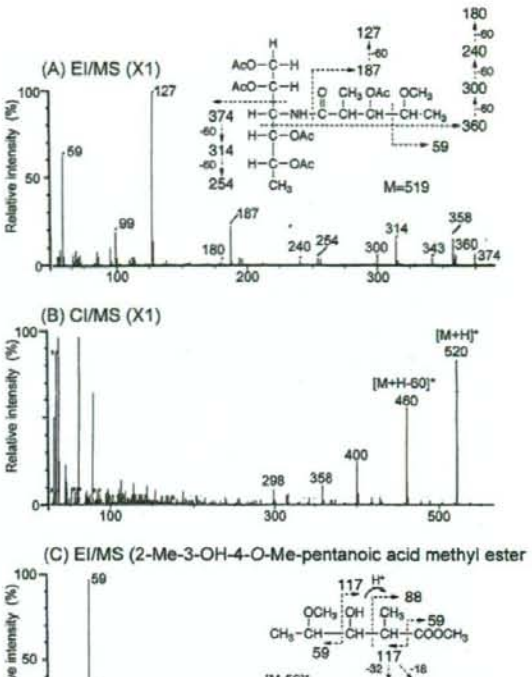

FIG. 3. EI-MS and CI-MS spectra of the alditol acetate derivative from X1 (A and B) and *N*-acylated-short-chain fatty acid methyl ester (C). The pattern of prominent fragment ions is illustrated. The CG column and condition were described in the legend for Fig. 2.

were performed. As shown in Fig. 4, the GC-MS spectra of perdeuteromethylated alditol acetates were assigned four major peaks, 1,3,4,5-tetra-*O*-deuteromethyl-2-*O*-acetyl-6-deoxy-talitol (*m/z* 109, 132, 154, 167, and 214); 2,4-di-*O*-deuteromethyl-1,3,5-tri-*O*-acetyl-rhamnitol (*m/z* 121, 134, 205, 240, and 253); 2-*O*-deuteromethyl-4-*O*-methyl-1,3,5-tri-*O*-acetyl-rhamnitol (*m/z* 121, 131, 202, and 237); and 2,4-di-*O*-deuteromethyl-1,5-di-*O*-acetyl-3-2′-methyl-3′-*O*-deuteromethyl-4′methoxy-pentanoyl-deuteromethylamido-3,6-dideoxy-hexitol (*m/z* 121, 134, and 341). These results revealed that the 6-d-Tal residue was linked at C-2; Rha and 4-*O*-Me-Rha were linked at C-1 and C-3; and the nonreducing terminus, 3-2′-methyl-3′-hydroxy-4′-methoxy-pentanoyl-amido-3,6-dideoxy-Hex, was C-1 linked. The MALDI-TOF/TOF MS spectrum of the oligoglycosyl alditol from serotype 16 OSE afforded the expected molecular ions $[M+Na]^+$ at *m/z* 931, together with the characteristic mass increments in the series of glycosyloxonium ions formed on fragmentation at *m/z* 312, 472, 618, and 764 from the terminal sugar *N*-acyl-Hex to 6-d-Tal and at *m/z* 336, 482, and 642 from 6-d-Tal to *N*-acyl-Hex (Fig. 5). Rha residues were determined to be in the L absolute configuration by comparative GC-MS analyses of trimethylsilylated (*S*)-(+)-sec-butyl glycosides and (*R*)-(−)-sec-butylglycosides (see Fig. S1 in the supplemental material). Taken together, these results established the sequence and linkage arrangement 3-2′-methyl-3′-hydroxy-4′-methoxy-pentanoyl-amido-3,6-dideoxy-Hex-(1→3)-4-*O*-Me-Rha-(1→3)-L-Rha-(1→3)-L-Rha-(1→2)-6-d-Tal, exclusively.

NMR analysis of serotype 16 OSE. The ^{1}H NMR and ^{1}H-^{1}H COSY analyses of the serotype 16 GPL revealed six distinct anomeric protons with corresponding H1-H2 cross peaks in the low field region at δ4.93, 4.92, 4.92, 4.84, 4.65 ($J_{1,2}$ = 2 to 3 Hz, indicative of α-anomers) and 4.51 (a doublet, $J_{1,2}$ = 7.7 Hz, indicative of a β-hexosyl unit). When further analyzed by ^{1}H-detected [^{1}H, ^{13}C] two-dimensional HMQC, the anomeric protons resonating at δ4.93, 4.92, 4.92, 4.84, 4.65, and 4.51 have C-1s resonating at δ101.57, 95.73, 101.40, 102.56, 100.97, and 103.36, respectively (for a summary, see Table S1 in the supplemental material). The J_{C1} values for each of these protons were calculated to be 171, 170, 171, 170, 169, and 161 Hz by measurement of the inverse-detection nondecoupled two-dimensional HMQC (Fig. 6). These results established that the terminal amido-Hex was a β configuration and the others were α-anomers.

Cloning and sequence of serotype 16 GPL biosynthesis cluster. To isolate the serotype 16 GPL biosynthesis cluster, the genomic cosmid library of *M. intracellulare* serotype 16 strain ATCC 13950^T was constructed. Primers were designed to amplify the region corresponding to the *rtfA* gene. More than 300 cosmid clones were tested using colony PCR with *rtfA* primers, and the positive clones no. 51 and 253 were isolated from the *E. coli* transductants. PCR analysis revealed that clone no. 253 contained a *drrC* gene but that clone no. 51 did not. Thus, we used clone no. 253 for subsequent sequence analysis for the *gtfB-drrC* region. The 22.9-kb region of *M. intracellulare* sero-

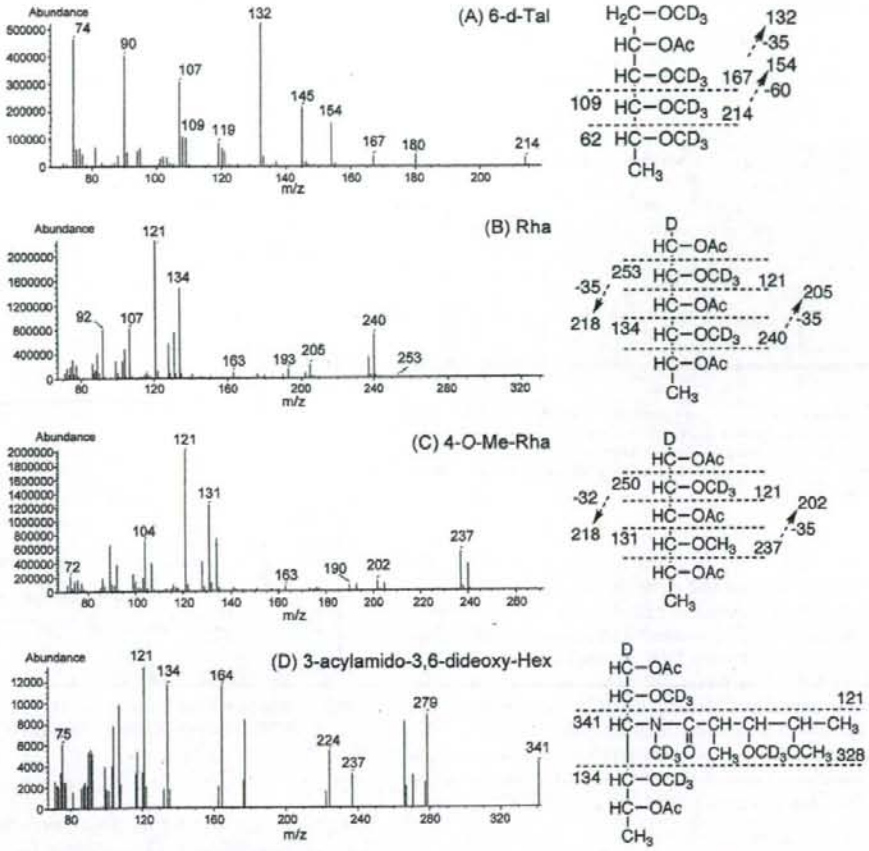

FIG. 4. GC-MS spectra of individual perdeuteromethylated alditol acetate derivatives from serotype 16 OSE. The formation of prominent fragment ions is illustrated; fragments were assigned to 1,3,4,5-tetra-*O*-deuteromethyl-2-*O*-acetyl-6-deoxy-talitol (A), 2,4-di-*O*-deuteromethyl-1,3,5-tri-*O*-acetyl-rhamnitol (B), 2-*O*-deuteromethyl-4-*O*-methyl-1,3,5-tri-*O*-acetyl-rhamnitol (C), and 2,4-di-*O*-deuteromethyl-1,5-di-*O*-acetyl-3-2′-methyl-3′-*O*-deuteromethyl-4′-methoxy-pentanoyl-deuteromethylamido-3,6-dideoxy-hexitol (D).

type 16 ATCC 13950^T was deposited in the NCBI GenBank database (accession no. AB355138). The similarity to protein sequences of each ORF is summarized in Table 1, and the genetic map for the serotype 16 GPL biosynthetic cluster was compared with those of serotype 2, 4, and 7 GPLs (Fig. 7). The *gtfB* and *drrC* genes of *M. intracellulare* serotype 16 ATCC 13950^T had 99.8% and 83.7% DNA identities with those of *M. intracellulare* serotype 7 ATCC 35847, respectively. In the DNA region between *gtfB* and *drrC* (20.8 kb), 17 ORFs were observed. Four ORFs (ORF 1, 2, 16, and 17) were homologous to those found in the same region of serotype 7-specific DNA, and the others were unique to the serotype 16 strain. No insertion of insertion elements or transposons was detected in this region. The nucleotide sequences of the ORF 1 and ORF 2 in serotype 16 strain ATCC 13950^T were homologous to those of ORF 1 and ORF 8 in serotype 7, respectively, suggesting that these two ORFs have the same function. The similarity of the deduced amino acid sequences suggested the possibility that the functions of ORF 3 and ORF 6 are to encode methyltransferase and aminotransferase, respectively. The deduced amino acid sequences of ORF 4 and ORF 5 showed significant similarities to the WxcM protein, the function of which is not clear. Interestingly, the deduced amino acid sequences of ORF 16 and ORF 17 of serotype 16 were homologous to ORF 9 of serotype 7. ORFs 1, 16, and 17 have considerable homology to glycosyltransferases. Nine ORFs, which are possibly involved in fatty acid synthesis, were detected between ORF 7 and ORF 15. It is notable that ORF 13 had a chimeric structure. The N-terminal half of ORF 13 showed similarity to phosphate butyryl/acetyl transferases, but the C-terminal half showed similarity to short-chain reductase/dehydrogenases. These results suggest that this region of DNA is responsible for the biosynthesis of the serotype 16-specific GPL.

Expression of cosmid clone no. 253 in *M. avium* serotype 1 strain. The OSE of serotype 1 GPL was composed of α-L-Rha-

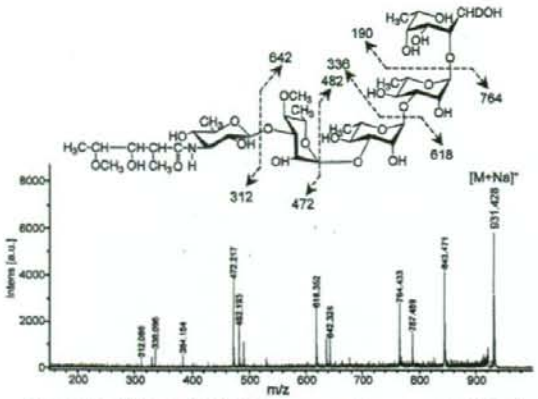

FIG. 5. MALDI-TOF/TOF MS spectrum of serotype 16 OSE. The formation of a characteristic increment in fragment ions is illustrated. The matrix was 10 mg/ml 2,5-dihydroxybenzoic acid in ethanol-water (3:7, vol/vol), and it was performed in the lift-lift mode. Intens., intensity; a.u., absorbance units.

(1→2)-6-d-L-Tal (9). The *M. avium* serotype 1 strain (NF113) was transformed with cosmid clone no. 253 containing a serotype 16-specific gene cluster and produced a new GPL with a different R_f value by TLC compared to serotype 1 GPL (Fig. 8A). The R_f value of the new GPL was identical to that of the serotype 16 GPL. The molecular weight of intact GPL, the fragment pattern of its OSE, and the GC pattern of the alditol acetate derivatives were completely equivalent to those of the serotype 16 GPL (see Fig. S2 in the supplemental material). As a result, the transformant of the serotype 1 strain expressed the cosmid clone no. 253 gene cluster and produced the serotype 16 GPL.

DISCUSSION

MAC species have serotype-specific GPLs that are characteristic components of the outer layer of the cell wall (6, 9). In addition to their serological differentiation, the chemical structures of 15 serotype-specific GPLs derived from the predominant clinical isolates have been analyzed; however, those of other GPLs remain unclear. The present study demonstrates the chemical structure of the serotype 16 GPL derived from *M. intracellulare*. We determined the glycosyl composition, linkage positions, and anomeric and ring configurations of the glycosyl residues of the serotype 16 GPL, and its OSE was defined as 3-2′-methyl-3′-hydroxy-4′-methoxy-pentanoyl-amido-3,6-dideoxy-β-Hex-(1→3)-4-*O*-methyl-α-L-Rha-(1→3)-α-L-Rha-(1→3)-α-L-Rha-(1→2)-6-d-α-L-Tal (Fig. 8B). The serotype 16 GPL should be listed as a group 2 polar GPL in the structural classification of Chatterjee and Khoo (9).

The GPLs of serotypes 7, 12, 17, and 19 have already been classified as group 2 GPLs, which are commonly composed of R→α-L-Rha-(1→3)-α-L-Rha-(1→2)-6-d-L-Tal (R, variable region), possessing a characteristic terminal sugar such as *N*-acyldeoxy-Hex. Indeed, the presence of an amido sugar has been reported in only five GPLs, serotypes 7, 12, 14, 17, and 25 (8, 9, 18). It has been determined that the OSE structure of the

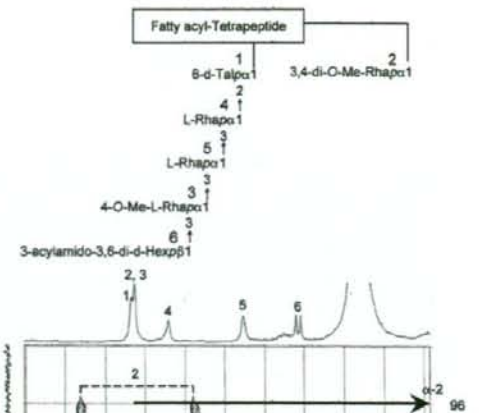

FIG. 6. Nondecoupled ^{1}H-detected [^{1}H, ^{13}C] HMQC spectrum of serotype 16 GPL. Cross-peak labels correspond to those shown on the structure.

serotype 17 GPL was 3-2′-methyl-3′-hydroxy-butanoyl-amido-3,6-dideoxy-β-D-Glc-(1→3)-4-*O*-methyl-α-L-Rha-(1→3)-α-L-Rha-(1→3)-α-L-Rha-(1→2)-6-d-L-Tal (9, 25). Based on the behavior of GPLs in TLC and the GC-MS analysis of alditol acetate derivatives, serotype 16 GPL seems to possess a unique carbohydrate epitope similar to that of serotype 17 GPL. We compared the OSE of serotype 16 GPL to that of serotype 17 GPL. The acylated amido group that was bound to the terminal sugar was different, although the linkage position was identical. Except for the terminal-acylated amido sugar, the other sugar compositions and glycosyl linkage positions were completely identical. An acylated amido group attached to the C-3 position of Hex is very unusual. To our knowledge, 3-amido-Hex is irregular in nature, although 2-amido-Hex is known to be glucosamine or galactosamine, which is frequently isolated as a component of lipopolysaccharides and glycosaminoglycans in prokaryotic and eukaryotic cells (7, 42). Further, existence of short-chain fatty acid 2-methyl-3-hydroxy-4-methoxy-pentanoic acid linked to the amido group of d-Hex is also unique. The characteristic gene cluster is thought to regulate the production of 3-acylated-amido-Hex. It is difficult to determine the species of acylated amido sugars, because no reference standard is available. The terminal sugar of the serotype 17 GPL was reviewed as a gluco-configuration, although firm evidence was not shown (9, 25). The J_{CH} and $J_{1\text{-}2}$ values for the anomeric proton in the terminal sugar were 161 and 7.7 Hz,

TABLE 1. Similarity to protein sequences of ORFs in cosmid clone no. 253 derived from *M. intracellulare* serotype 16 strain ATCC 13950^T

ORF	Predicted molecular mass (kDa)	Predicted pI	Exhibits similarity to:	E value	Amino acid identity (no. matched/total no.)	Accession no.
GtfB	45.6	6.35	Glycosyltransferase GtfB	0.0	417/418	BAF45360
Orf 1	45.2	6.10	Putative glycosyltransferase	0.0	416/417	BAF45361
Orf 2	78.9	8.51	Putative acyltransferase	0.0	557/728	BAF45368
Orf 3	31.0	5.88	Putative methyltransferase	2e-89	382/421	NP_218045
Orf 4	15.7	4.94	Conserved hypothetical protein	1e-39	73/129	BAD50406
Orf 5	16.0	4.69	Conserved hypothetical protein	5e-40	75/135	EAX55190
Orf 6	41.1	5.88	Aminotransferase/DegT_DnrJ_EryC1	6e-119	208/357	ABD68440
Orf 7	40.6	9.65	Conserved hypothetical protein	2e-89	178/304	AAS03547
Orf 8	36.7	5.32	Conserved hypothetical protein	2e-52	116/298	CAE06954
Orf 9	22.3	9.79	Putative *N*-acetyltransferase	4e-14	58/166	EAU11841
Orf 10	25.3	7.82	Short-chain dehydrogenase/reductase	7e-47	101/233	EAO61220
Orf 11	23.8	6.05	Putative hydrolase	4e-24	64/196	ABG85599
Orf 12	37.2	6.50	Ketoacyl-acyl carrier protein synthase III	3e-55	126/331	EAX48715
Orf 13	42.5	7.72	Short-chain dehydrogenase/reductase	2e-42	97/248	ZP_01289005
Orf 14	65.8	4.70	Predicted enzyme involved in methoxymalonyl-acyl carrier protein biosynthesis	6e-85	201/575	ABB73590
Orf 15	50.0	6.23	Acyl coenzyme A synthetases	2e-128	233/445	EAT27362
Orf 16	39.1	8.00	Putative glycosyltransferase	2e-106	196/318	NP_855197
Orf 17	37.7	9.46	Putative glycosyltransferase	8e-160	278/323	BAF45369
DrrC	28.6	11.47	Daunorubicin resistance protein C	6e-132	233/261	BAF45370

respectively (Fig. 6; Table S1 in the supplemental material). These results demonstrated unequivocally that the terminal amido-Hex was β configuration and H-2 was in the axial position. The terminal amido-Hex is considered to be derived from glucose or galactose, not Rha.

Next, we explored the genetic mechanism of GPL biosynthesis, because the elongation of carbohydrate chains in serotype-specific GPLs is poorly understood. The *ser2* gene cluster of the *M. avium* serotype 2 strain (31) and a 27.5-kb DNA fragment of the *M. avium* serotype 4 strain (28) were identified to be responsible for the biosynthesis of each OSE in GPLs. Recently, enzymatic characterizations of glycosyltransferase and methyltransferase of nonpolar GPLs have been reported for *Mycobacterium smegmatis* (36, 38). In the serotype-specific polar GPL biosynthesis of MAC, only the *rtfA* gene was functionally clarified to encode the transfer of L-Rha to 6-d-Tal, but which gene cluster transfers the sugars next to L-Rha elongated from 6-d-Tal is unclear.

In this study, we cloned the biosynthetic cluster of the serotype 16 GPL and analyzed its sequence. Seventeen ORFs were detected in the serotype 16 strain, and the sequence homology was analyzed. The transformant of the *M. avium* serotype 1

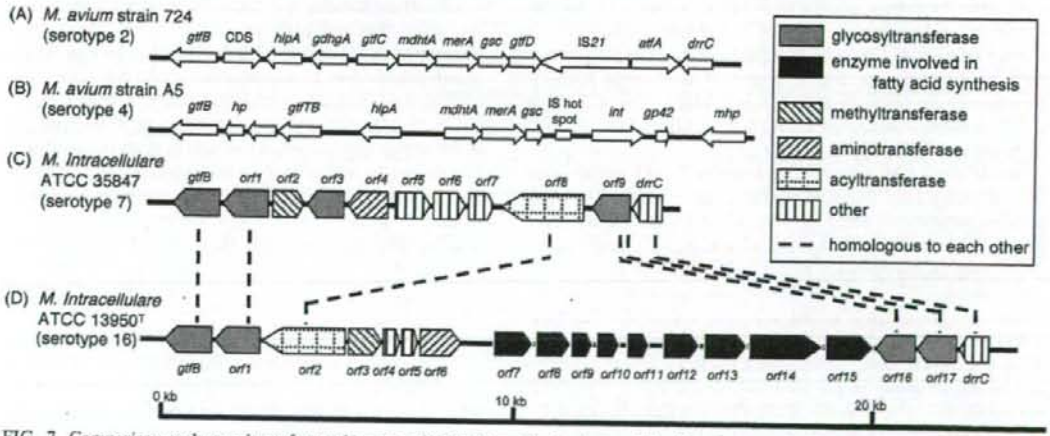

FIG. 7. Comparison and overview of genetic maps of GPL biosynthetic cluster. The *M. avium* strain 724 annotated sequence obtained from GenBank (accession no. AF125999) (A); the *M. avium* strain A5 annotated sequence obtained from GenBank (accession no. AY130970) (B); the *M. intracellulare* ATCC 35847 sequenced in our previous study (GenBank accession no. AB274811) (C); the *M. intracellulare* ATCC 13950^T sequenced in this study (GenBank accession no. AB355138) (D). The orientation of each gene is shown by the direction of the arrow. In panels A and B, putative ORFs not showing homology to known proteins sequences are not depicted. The sequences extending upstream in panels A and B and downstream in panel B are not included in the figure.

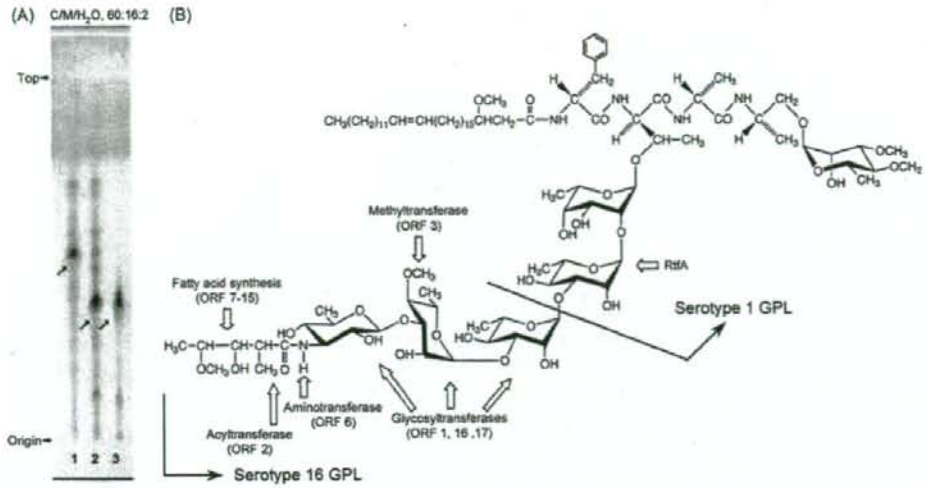

FIG. 8. TLC pattern of *M. avium* serotype 1 and its transformant with cosmid clone no. 253 and proposed complete structure of the serotype 16 GPL. (A) The alkaline-stable lipids derived from *M. avium* serotype 1 (lane 1), its transformant (lane 2), and purified serotype 16 GPL (lane 3) were developed with the solvent system of chloroform-methanol-water (60:16:2, vol/vol/vol). (B) Predicted biosynthesis gene clusters are indicated by arrows.

strain carrying cosmid clone no. 253 produced serotype 16 GPL. These results strongly implied that this *gtfB-drrC* region is responsible for the biosynthesis of the serotype 16-specific GPL. From the structural analysis of the serotype 16 GPL and the sequence of cosmid clone no. 253, it is possible to predict the relationship between the biosynthesis of serotype 16 GPL and the function of each ORF.

The genetic map of the serotype 16 GPL biosynthetic cluster was compared to those of serotype 2 GPL from *M. avium* strain 724, serotype 4 GPL from *M. avium* strain A5, and serotype 7 GPL from *M. intracellulare* strain ATCC 35847^T (12, 18, 28). Significant differences were found in the neighborhood of the conserved region. The genetic organization of the serotype 16 GPL gene cluster was distinct from that of serotype 7, except for some of the ORFs, and the ORFs in this region of serotype 2 and serotype 4 were completely different from ORFs 1 to 17 in serotype 16 (Fig. 7).

In addition to *M. intracellulare* serotype 7 (18) and serotype 16 strains, we have analyzed similar gene clusters of *M. intracellulare* serotype 12 and 17 strains. The sequence homology of the regions of ORF 1 and ORF 17 was highly conserved between only *M. intracellulare* serotype 16 and 17 strains (unpublished data). ORFs 1, 16, and 17 may lead to transfer of the two additional molecules of L-Rha and terminal amido-Hex. ORF 2 was assigned to acyltransferase and may be responsible for biosynthesis of the 3-2′-methyl-3′-hydroxy-4′-methoxy-pentanoyl-amido group in the terminal Hex. ORF 3 is probably responsible for the transfer of the *O*-methyl group at the C-4 position in the third L-Rha from 6-d-Tal. ORF 6 is homologous to aminotransferase and possibly associated with the biosynthesis of an amido group in the terminal Hex. The deduced amino acid sequences of ORF 6 in serotype 16 and ORF 4 in serotype 7 have homologies to DegT_DnrJ_EryC1 aminotransferases. However, these two ORFs are dissimilar to each other. Serotype 16 and 7 GPLs have an amido group at the terminal Hex, although the attachment position is different. The serotype 7 GPL has an amido group at the C-4 position in the terminal Hex, but the serotype 16 GPL has it at the C-3 position. Nine ORFs between ORF 7 and ORF 15 are possibly involved in fatty acid synthesis of the acyl chain moiety linked by an amido bond of the terminal Hex. Taken together, this gene cluster may participate in the biosynthetic pathway of the serotype 16-specific GPL, although further study is needed to clarify the function of each ORF.

Recent studies suggest that GPLs play an important role in the phenotype and pathogenicity of MAC. The colony morphology is considered to be influenced by cell wall GPL. MAC colony phenotypes spontaneously occur from smooth to rough type, and this is due to a mutation lacking GPL (3, 13, 22). The deletion of genomic regions encoding GPL biosynthesis may result in the loss of GPL. Danelishvili et al. demonstrated that the uptake by and growth in macrophages of a MAC mutant with the gene belonging to the GPL synthesis pathway inactivated by transposon insertion were decreased (11). Bhatnagar and Schorey have reported that macrophages infected with MAC release exosomes containing GPLs that result in the transfer of the GPLs to uninfected macrophages and induce a proinflammatory response (4). These findings imply that GPL participates in the pathogenicity of MAC. By contrast, our previous studies have demonstrated that anti-GPL antibodies are detected in the sera of most immunocompetent patients with MAC pulmonary disease and that the detection of anti-GPL antibody is useful for the serodiagnosis of MAC disease (15, 26, 27).

To understand the role of GPLs in MAC and its hosts, it is necessary to define the chemical structure and biosynthesis pathways of GPLs. Elucidation of the structure-function relationship of GPL may open a new avenue for controlling MAC disease.

ACKNOWLEDGMENTS

This work was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, the Japan Health Sciences Foundation, and the Ministry of Health, Labor, and Welfare of Japan (Research on Emerging and Reemerging Infectious Diseases).

We are grateful to Sumihiro Hase (Department of Chemistry, Graduate School of Science, Osaka University, Osaka, Japan) and Hiromi Murakami (Osaka Municipal Technical Research Institute, Osaka, Japan) for helpful discussion.

REFERENCES

1. **Baess, I.** 1983. Deoxyribonucleic acid relationships between different serovars of *Mycobacterium avium*, *Mycobacterium intracellulare* and *Mycobacterium scrofulaceum*. Acta Pathol. Microbiol. Immunol. Scand. **91**:201–203.
2. **Barrow, W. W., T. L. Davis, E. L. Wright, V. Labrousse, M. Bachelet, and N. Rastogi.** 1995. Immunomodulatory spectrum of lipids associated with *Mycobacterium avium* serovar 8. Infect. Immun. **63**:126–133.
3. **Belisle, J. T., K. Klaczkiewicz, P. J. Brennan, W. R. Jacobs, Jr., and J. M. Inamine.** 1993. Rough morphological variants of *Mycobacterium avium*. Characterization of genomic deletions resulting in the loss of glycopeptidolipid expression. J. Biol. Chem. **268**:10517–10523.
4. **Bhatnagar, S., and J. S. Schorey.** 2007. Exosomes released from infected macrophages contain *Mycobacterium avium* glycopeptidolipids and are proinflammatory. J. Biol. Chem. **282**:25779–25789.
5. **Bhatt, A., N. Fujiwara, K. Bhatt, S. S. Gurcha, L. Kremer, B. Chen, J. Chan, S. A. Porcelli, K. Kobayashi, G. S. Besra, and W. R. Jacobs, Jr.** 2007. Deletion of *kasB* in *Mycobacterium tuberculosis* causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. Proc. Natl. Acad. Sci. USA **104**:5157–5162.
6. **Brennan, P. J., and H. Nikaido.** 1995. The envelope of mycobacteria. Annu. Rev. Biochem. **64**:29–63.
7. **Campo, G. M., S. Campo, A. M. Ferlazzo, R. Vinci, and A. Calatroni.** 2001. Improved high-performance liquid chromatographic method to estimate aminosugars and its application to glycosaminoglycan determination in plasma and serum. J. Chromatogr. B **765**:151–160.
8. **Chatterjee, D., G. O. Aspinall, and P. J. Brennan.** 1987. The presence of novel glucuronic acid-containing, type-specific glycolipid antigens within *Mycobacterium* spp. Revision of earlier structures. J. Biol. Chem. **262**:3528–3533.
9. **Chatterjee, D., and K. H. Khoo.** 2001. The surface glycopeptidolipids of mycobacteria: structures and biological properties. Cell. Mol. Life Sci. **58**: 2018–2042.
10. **Daffe, M., and P. Draper.** 1998. The envelope layers of mycobacteria with reference to their pathogenicity. Adv. Microb. Physiol. **39**:131–203.
11. **Danelishvili, L., M. Wu, B. Stang, M. Harriff, S. Cirillo, J. Cirillo, R. Bildfell, B. Arbogast, and L. E. Bermudez.** 2007. Identification of *Mycobacterium avium* pathogenicity island important for macrophage and amoeba infection. Proc. Natl. Acad. Sci. USA **104**:11038–11043.
12. **Eckstein, T. M., J. T. Belisle, and J. M. Inamine.** 2003. Proposed pathway for the biosynthesis of serovar-specific glycopeptidolipids in *Mycobacterium avium* serovar 2. Microbiology **149**:2797–2807.
13. **Eckstein, T. M., J. M. Inamine, M. L. Lambert, and J. T. Belisle.** 2000. A genetic mechanism for deletion of the *ser2* gene cluster and formation of rough morphological variants of *Mycobacterium avium*. J. Bacteriol. **182**: 6177–6182.
14. **Eckstein, T. M., F. S. Silbaq, D. Chatterjee, N. J. Kelly, P. J. Brennan, and J. T. Belisle.** 1998. Identification and recombinant expression of a *Mycobacterium avium* rhamnosyltransferase gene (*rtfA*) involved in glycopeptidolipid biosynthesis. J. Bacteriol. **180**:5567–5573.
15. **Enomoto, K., S. Oka, N. Fujiwara, T. Okamoto, Y. Okuda, R. Maekura, T. Kuroki, and I. Yano.** 1998. Rapid serodiagnosis of *Mycobacterium avium-intracellulare* complex infection by ELISA with cord factor (trehalose 6, 6′-dimycolate), and serotyping using the glycopeptidolipid antigen. Microbiol. Immunol. **42**:689–696.
16. **Falkinham, J. O., III.** 1996. Epidemiology of infection by nontuberculous mycobacteria. Clin. Microbiol. Rev. **9**:177–215.
17. **Field, S. K., D. Fisher, and R. L. Cowie.** 2004. *Mycobacterium avium* complex pulmonary disease in patients without HIV infection. Chest **126**:566–581.
18. **Fujiwara, N., N. Nakata, S. Maeda, T. Naka, M. Doe, I. Yano, and K. Kobayashi.** 2007. Structural characterization of a specific glycopeptidolipid containing a novel *N*-acyl-deoxy sugar from *Mycobacterium intracellulare* serotype 7 and genetic analysis of its glycosylation pathway. J. Bacteriol. **189**:1099–1108.
19. **Gerwig, G. J., J. P. Kamerling, and J. F. G. Vliegenthart.** 1978. Determination of the D and L configuration of neutral monosaccharides by high-resolution capillary G.L.C. Carbohydr. Res. **62**:349–357.
20. **Hakomori, S.** 1964. A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. J. Biochem. (Tokyo) **55**:205–208.
21. **Heidelberg, T., and O. R. Martin.** 2004. Synthesis of the glycopeptidolipid of *Mycobacterium avium* serovar 4: first example of a fully synthetic C-mycoside GPL. J. Org. Chem. **69**:2290–2301.
22. **Howard, S. T., E. Rhoades, J. Recht, X. Pang, A. Alsup, R. Kolter, C. R. Lyons, and T. F. Byrd.** 2006. Spontaneous reversion of *Mycobacterium abscessus* from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype. Microbiology **152**:1581–1590.
23. **Kaufmann, S. H.** 2001. How can immunology contribute to the control of tuberculosis? Nat. Rev. Immunol. **1**:20–30.
24. **Khoo, K. H., D. Chatterjee, A. Dell, H. R. Morris, P. J. Brennan, and P. Draper.** 1996. Novel *O*-methylated terminal glucuronic acid characterizes the polar-glycopeptidolipids of *Mycobacterium habana* strain TMC 5135. J. Biol. Chem. **271**:12333–12342.
25. **Khoo, K. H., E. Jarboe, A. Barker, J. Torrelles, C. W. Kuo, and D. Chatterjee.** 1999. Altered expression profile of the surface glycopeptidolipids in drug-resistant clinical isolates of *Mycobacterium avium* complex. J. Biol. Chem. **274**:9778–9785.
26. **Kitada, S., R. Maekura, N. Toyoshima, N. Fujiwara, I. Yano, T. Ogura, M. Ito, and K. Kobayashi.** 2002. Serodiagnosis of pulmonary disease due to *Mycobacterium avium* complex with an enzyme immunoassay that uses a mixture of glycopeptidolipid antigens. Clin. Infect. Dis. **35**:1328–1335.
27. **Kitada, S., Y. Nishiuchi, T. Hiraga, N. Naka, H. Hashimoto, K. Yoshimura, K. Miki, M. Miki, M. Motone, T. Fujikawa, K. Kobayashi, I. Yano, and R. Maekura.** 2007. Serological test and chest computed tomography findings in patients with *Mycobacterium avium* complex lung disease. Eur. Respir. J. **29**:1217–1223.
28. **Krzywinska, E., and J. S. Schorey.** 2003. Characterization of genetic differences between *Mycobacterium avium* subsp. avium strains of diverse virulence with a focus on the glycopeptidolipid biosynthesis cluster. Vet. Microbiol. **91**:249–264.
29. **Maekura, R., Y. Okuda, A. Hirotani, S. Kitada, T. Hiraga, K. Yoshimura, I. Yano, K. Kobayashi, and M. Ito.** 2005. Clinical and prognostic importance of serotyping *Mycobacterium avium-Mycobacterium intracellulare* complex isolates in human immunodeficiency virus-negative patients. J. Clin. Microbiol. **43**:3150–3158.
30. **Marras, T. K., and C. L. Daley.** 2002. Epidemiology of human pulmonary infection with nontuberculous mycobacteria. Clin. Chest Med. **23**:553–567.
31. **Maslow, J. N., V. R. Irani, S. H. Lee, T. M. Eckstein, J. M. Inamine, and J. T. Belisle.** 2003. Biosynthetic specificity of the rhamnosyltransferase gene of *Mycobacterium avium* serovar 2 as determined by allelic exchange mutagenesis. Microbiology **149**:3193–3202.
32. **McClatchy, J. K.** 1981. The seroagglutination test in the study of nontuberculous mycobacteria. Rev. Infect. Dis. **3**:867–870.
33. **McCloskey, J. A.** 1969. Mass spectrometry, p. 402. *In* J. M. Lowenstein (ed.), Methods in enzymology: lipid, vol. 14. Academic Press, New York, NY.
34. **McNeil, M., H. Gaylord, and P. J. Brennan.** 1988. *N*-formylkansosaminyl-(1-3)-2-*O*-methyl-D-rhamnopyranose: the type-specific determinant of serovariant 14 of the *Mycobacterium avium* complex. Carbohydr. Res. **177**:185–198.
35. **McNeil, M., A. Y. Tsang, and P. J. Brennan.** 1987. Structure and antigenicity of the specific oligosaccharide hapten from the glycopeptidolipid antigen of *Mycobacterium avium* serotype 4, the dominant *Mycobacterium* isolated from patients with acquired immune deficiency syndrome. J. Biol. Chem. **262**: 2630–2635.
36. **Miyamoto, Y., T. Mukai, N. Nakata, Y. Maeda, M. Kai, T. Naka, I. Yano, and M. Makino.** 2006. Identification and characterization of the genes involved in glycosylation pathways of mycobacterial glycopeptidolipid biosynthesis. J. Bacteriol. **188**:86–95.
37. **Odham, G., and E. Stenhagen.** 1972. Fatty acids, p. 211–228. *In* G. R. Waller (ed.), Biochemical application of mass spectrometry. Wiley-Interscience, New York, NY.
38. **Patterson, J. H., M. J. McConville, R. E. Haites, R. L. Coppel, and H. Billman-Jacobe.** 2000. Identification of a methyltransferase from *Mycobacterium smegmatis* involved in glycopeptidolipid synthesis. J. Biol. Chem. **275**:24900–24906.
39. **Supply, P., E. Mazars, S. Lesjean, V. Vincent, B. Gicquel, and C. Locht.** 2000. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol. Microbiol. **36**:762–771.
40. **Tsang, A. Y., J. C. Denner, P. J. Brennan, and J. K. McClatchy.** 1992. Clinical and epidemiological importance of typing of *Mycobacterium avium* complex isolates. J. Clin. Microbiol. **30**:479–484.
41. **Wayne, L. G., and H. A. Sramek.** 1992. Agents of newly recognized or infrequently encountered mycobacterial diseases. Clin. Microbiol. Rev. **5**:1–25.
42. **Woods, A., and J. R. Couchman.** 2001. Proteoglycan isolation and analysis, p. 10.7.1–10.7.19. *In* J. S. Bonifacino, M. Dasso, J. B. Harford, J. Lippincott-Schwartz, and K. M. Yamada (ed.), Current protocols in cell biology. Wiley Interscience, Hoboken, NJ.

[Prion 2:2, 73-80; April/May/June 2008]; ©2008 Landes Bioscience

Research Paper

Anti-bovine prion protein RNA aptamer containing tandem GGA repeat interacts both with recombinant bovine prion protein and its β isoform with high affinity

Kazuyoshi Murakami,[1,2] Fumiko Nishikawa,[2] Ken Noda,[3] Takashi Yokoyama[4] and Satoshi Nishikawa[1,2,*]

[1]Graduate School of Life and Environmental Science; University of Tsukuba; Tennodai; Tsukuba, Ibaraki Japan; [2]Age Dimension Research Center; National Institute of Advanced Industrial Science and Technology (AIST); Tsukuba, Ibaraki Japan; [3]National Veterinary Assay Laboratory (NVAL); Ministry of Agriculture; Forestry and Fisheries; Kokubunji, Tokyo Japan; [4]Prion Disease Research Center; National Institute of Animal Health (NIAH); Tsukuba, Ibaraki Japan

Abbreviations: bPrP, bovine prion protein; bPrP-β, β isoform of bovine prion protein; bPrP (102–241), amino acids 102–241 of bovine prion protein; bPrP (132–241), amino acids 132–241 of bovine prion protein; mPrP, mouse prion protein; PrP^{C}, cellular isoform of prion protein; PrP^{Sc}, scrapie isoform of prion protein; SELEX, systematic evolution of ligands by exponential enrichment; CD, circular dichroism

Key words: RNA aptamer, prion protein, SELEX, GGA repeat, G-quadruplex

In order to obtain RNA aptamers against bovine prion protein (bPrP), we carried out in vitro selection from RNA pools containing a 55-nucleotide randomized region to target recombinant bPrP. Most of obtained aptamers contained conserved GGA tandem repeats $(GGA)_4$ and aptamer #1 (apt #1) showed a high affinity for both bPrP and its β isoform (bPrP-β). The sequence of apt #1 suggested that it would have a G-quadruplex structure, which was confirmed using CD spectra in titration with KCl. A mutagenic study of this conserved region, and competitive assays, showed that the conserved $(GGA)_4$ sequence is important for specific binding to bPrP and bPrP-β. Following 5'-biotinylation, aptamer #1 specifically detected PrP^{c} in bovine brain homogenate in a Northwestern blotting assay. Protein deletion mutant analysis showed that the bPrP aptamer binds within 25–131 of the bPrP sequence. Interestingly, the minimized aptamer #1 (17 nt) showed greater binding to bPrP and bPrP-β as compared to apt #1. This minimized form of aptamer #1 may therefore be useful in the detection of bPrP, diagnosis of prion disease, enrichment of bPr and ultimately in gaining a better understanding of prion diseases.

Introduction

The prion protein (PrP) has two alternative forms: a normal cellular protein (PrP^{C}), which is a soluble α-helix rich isoform, and an insoluble β-sheet rich abnormal isoform[1] known as the protease-resistant form (PrP^{Sc}).[2] The detailed mechanism of the structural conversion between the soluble and insoluble forms remains unknown. PrP^{C} is almost ubiquitously expressed and highly conserved in mammals, being anchored on the surface of cells. To gain a better insight into its putative role, much research has gone towards the discovery of molecules that bind prion protein. Of those identified, the most notable are divalent metal ions, several proteins and nucleic acids.[3] Knowledge of the function of PrP would contribute towards a better understanding of the processes involved in both the amplification of the infectious agent and the neuronal damage leading to the neuro-degeneration observed in prion diseases. As concerns about the transmission of prion disease in the fields of medicine and food safety increase, there is an increasing demand for an understanding of the processes by which PrP^{Sc} can be detected and addressed.[4,5]

Specific anti-PrP probes, which have high specificity and sensitivity, are required for the diagnosis of prion diseases. Clarification of the binding mechanism between PrP and such probes will assist structural study of PrP. Aptamers are artificial nucleotides derived from systematic evolution of ligands by exponential enrichment (SELEX),[6,7] which bind a wide range of targets with a high affinity and specificity, as is seen with antibodies.[8,9] Using this method, many aptamers specific to PrP^{10-20} and PrP^{Sc},[12,13] have been isolated.

Recently we have performed an in vitro selection for mouse PrP using an RNA pool with a 30 nt randomized region.[16] To prepare a repertoire of anti-PrP RNA aptamers, we carried out SELEX for recombinant bPrP and analyzed for bPrP and bPrP-β. The obtained major RNA aptamers bind to both bPrP and bPrP-β. They possessed tandem GGA repeats $(GGA)_4$ as a consensus sequence, and suggested to form a parallel G-quadruplex structure in the presence of K^{+} ions. We identified that this core sequence is important for specific binding to bovine PrPs and increased the affinity and selectivity by redesigning this core sequence.

Results

In vitro selection. To obtain RNA aptamers targeting recombinant bPrP, we carried out SELEX using 97-nt RNA pool that has

*Correspondence to: Satoshi Nishikawa; Age Dimension Research Center; National Institute of Advanced Industrial Science and Technology (AIST); Tsukuba, Ibaraki 305-8566 Japan; Tel.: +81.29.861.6097; Fax: +81.29.861.6095; Email: satoshi-nishikawa@aist.go.jp

Submitted: 08/25/08; Accepted: 09/17/08

Previously published online as a *Prion* E-publication: http://www.landesbioscience.com/journals/prion/article/7024

55-nt randomized sequences. To increase the selection stringency, we applied the following selection pressures (Table 1): (1) decreased protein concentration and reaction time, and increased washing volumes; (2) increased tRNA concentration as a non-specific competitor; and (3) increased concentration of anti-mPrP RNA aptamer[16] as a specific competitor. Furthermore we carried out mutagenic PCR[21] to introduce mutations into concentrated sequences (see Materials and Methods). As the enrichment of specific binding RNAs for bPrP was observed in the iterative process (data not shown), G10 RNA pool for bPrP was cloned and sequenced. Interestingly, all RNAs contained tandem GGA repeats, mainly four continuous GGA triplet repeats $(GGA)_4$ (shown underlined in Table 2).

Binding affinities of apt #1 against bPrP or bPrP-β. The representative RNA aptamers designated apt #1 and apt #6 were analyzed. To determine their respective K_d values, a filter binding assay was performed using 10 nM RNA and various concentrations of bPrP or bPrP-β in selection buffer [20 mM Tris-HCl (pH 7.5), 100 mM NaCl]. Apt #1 showed a higher affinity (Fig. 1, K_d = 82 ± 21 nM for bPrP) than that of apt #6 (data not shown, K_d = 166 ± 55 nM for bPrP). Because of the structural difference between bPrP and bPrP-β, binding affinity of both RNA aptamers for bPrP-β are over 10-fold lower than those for bPrP. As it demonstrated a higher affinity than apt #6, we focused on apt #1 in subsequent analyses.

The $(GGA)_4$ sequence suggested that it forms a G-quartet structure. The formation of a G-quartet structure is stabilized by monovalent cations such as K^+ and Na^+ ions.[25,26] In general, DNA/RNA G-quadruplexes bind K^+ over Na^+ ions.[25,26] As is clearly shown in Figure 1, binding affinities for both proteins bPrP and bPrP-β in 10 mM K^+ ion (K_d = 31 nM and 220 nM, respectively; solid lines) were increased in comparison with 100 mM Na^+ ion (dashed lines). We also tested apt #1 binding without salt (K_d = 50 nM for bPrP and ≥1000 nM for bPrP-β; data not shown). The presence of potassium ions conferred a greater affinity of apt #1 to bPrPs, which was particularly profound for bPrP-β. We therefore prepared a buffer solution containing K^+ ions to be used in subsequent analyses on apt #1 (20 mM Tris-HCl (pH 7.5), 10 mM KCl).

Detection of $bPrP^C$ in bovine brain homogenate using apt #1. To test the ability of apt #1 to detect $bPrP^C$, we performed Northwestern blotting using a 5'-biotinylated apt #1 (Bi#1) with streptavidin-alkaline phosphatase conjugates (SA-AP) as a secondary probe (left in Fig. 2A). Bi#1 detected the three types of $bPrP^C$ (non-, mono- and di-glycosylated forms of PrP^C; lane 1) in bovine brain homogenate as well as antibody T2 (lane 2). The epitope of antibody T2 is located between amino acids 147 and 152.[23] This result indicates that apt #1 is applicable as a detection tool for bPrP and may perform as well as antibodies.

Table 1 **In vitro selection conditions for recombinant bPrP**

Round of selection	RNA (μM)	bPrP (μM)	tRNA (μM)	Anti-PrP RNA aptamer[16] (μM)	Time (min)
1	10	2	0	0	60
2	6	0.4	10	0	30
3	4	0.2	40	8	20
4	3	0.1	100	16	15
5	2	0.3	100	32	10
6	2	0.05	100	48	10
7	2	0.05	*6	0	10
8**	2	0.05	100	64	10
9**	2	0.05	100	64	10
10	6	0.05	100	96	10

*U_{12} was used instead of tRNA. **The mutagenic PCR was introduced.

The detection limit of bPrP with Bi#1 was approximately 30 ng (spot 6 in the upper of Fig. 2B) using dot blot analysis in the presence of poly (U) as a non-competitive binder. A negative control BiC#1 that has the complementary sequence of apt #1, showed a detection limit 240 ng (spot 3 in the lower of Fig. 2B), indicating specific binding of apt #1 against $bPrP^C$.

Conserved region $(GGA)_4$ suggests a parallel G-quadruplex. To examine the structure of apt #1, we measured CD spectra of the aptamer in the different concentrations of K^+ ion. It is known that a parallel G-quadruplex gives a positive CD peak at 260 nm, while an antiparallel G-quadruplex gives a positive CD peak at 295 nm either with or without a positive CD peak at 260 nm in the presence of monovalent cations.[27,28] A peak was observed at 260 nm of CD and the peak intensity increased as the concentration of K^+ ions increased, with maximum peak intensity at 10 mM KCl (Fig. 3A). This result suggests that apt #1 forms a parallel G-quadruplex in the presence of K^+ ion, with saturation occurring at 10 mM KCl. We also confirmed no increase of peak intensity at 260 nm in the presence of K^+ ion on the apt #1 mutant with substitutions of G tracts in the conserved $(GGA)_4$ region (data not shown).

Matsugami et al. reported that $d(GGA)_4$ folds into an intramolecular G-quadruplex composed of a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad with parallel orientation under physiological K^+ conditions[24] (Fig. 3B). The CD spectra of apt #1 (Fig. 3A) showed a similar pattern with that of $d(GGA)_4$. Although direct evidence is lacking, we believe that apt #1 also forms the similar type of the parallel G-quadruplex structure. Previous reports

Table 2 **Sequences of randomized regions and majorities of isolated RNA aptamers against recombinant bPrP**

Name of clone	Number of clones	Sequence of randomized region
#1	14/30	5'-CAAUCCAUUCAUCUCUCGAAUGAGGAAGUAGCCCAAGAGGAGGAGGAGGAUGAGC-3'
#6	11/30	5'-ACCUUCUGUUCAUAUCGUGACCAACCCAAUAGAUUGUGAUAAAGGAGGAGGAGGA-3'
#20	3/30	5'-UUGCCAAUAGCCCUACGGAGGAGGAGGCUGGACUAUAGACAGUUUACUUAAUAAA-3'
others		5'-CGCUCUGCCUGCGACCCAUCGACGUGAGUUUGCGAAAAAAGGAGGAGGAGGA-3'
		5'-UAGCCCCCCUAGGAGCCGCUUGUGAUAAGAGUUCAAAUUGGAGGAGGUGGAUUGA-3'

The conserved GGA triplet repeats $(GGA)_4$ are underlined involving its one point mutations.

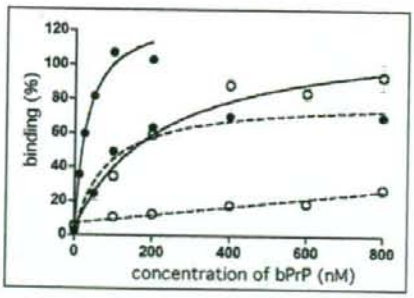

Figure 1. Binding of apt #1 to bPrP and bPrP-β. Binding curves of apt #1 to bPrP and bPrP-β are shown by closed circles and open circles, respectively. Solid and dashed lines represent different buffer conditions: 10 mM K^+ and 100 mM Na^+, respectively. The binding data are analyzed by GraphPad PRISM (see Materials and Methods).

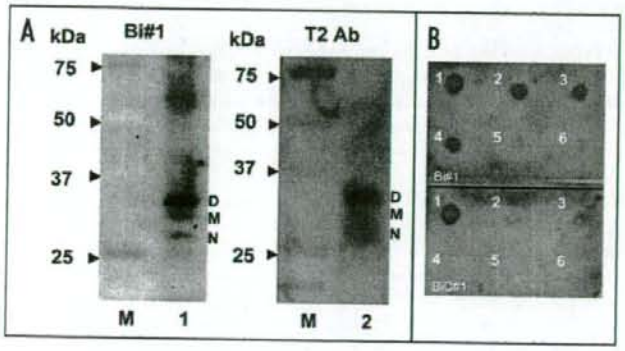

Figure 2. Detection of $bPrP^C$ in bovine brain homogenate and dot-blotting assay using a biotinylated apt #1 (Bi#1). (A) Bovine brain homogenate was analyzed using SDS-PAGE and subsequently with Northwestern blotting. $bPrP^C$ (lane 1 and 2) was detected using Bi#1 (left) and T2 antibody (right). Lane M represents protein molecular weight markers. N, M and D indicate non-, mono- and di-glycosilated PrP isoforms, respectively. (B) Dot-blotting assay of recombinant bPrP using Bi#1 or BiC#1. Number 1 represents cross-linked 5'-biotinylated RNA (Bi#1, BiC#1). Numbers 2–6 represent decreasing concentration of mounted bPrP (ng) as follows: 2 = 480, 3 = 240, 4 = 120, 5 = 60, 6 = 30.

have described G-quadruplex forming RNA aptamers for PrP,[10,11] but sequences are different from described here.

Aptamer binding sites locate amino acids 25–131. To investigate the apt #1 binding region of bPrP, we performed binding analysis with deleted proteins: bPrP (102–241) and bPrP (132–241) (Fig. 4A). Apt #1 bound strongly to full length bPrP (K_d = 31 nM), while K_d value of the deletion variant bPrP (102–241) (≥1000 nM; indicated with triangles) increased more than 30-fold compared with bPrP. Apt #1 could not bind to variant bPrP (132–241). Consequently, the binding region for apt #1 is located within amino acid residues 25–131, underlined in Figure 4B. Deletion of N-terminal 25–101 residues produced a strong effect for the aptamer binding. This suggests that this region is the main binding site as well as an anti-mPrP aptamer.[16] Mercey et al. reported that the two lysine clusters (amino acids 25–32 and 102–110 in ovine PrP) are primary nucleic acids binding sites, the former region showing stronger binding than the latter.[15] Considering these results, it seems that apt #1 is also likely to bind to the two lysine clusters, although there is no homology between apt #1 and ovine-PrP aptamers.

The conserved region $(GGA)_4$ is important for binding to bPrP or bPrP-β. To investigate the correlation between binding affinity and G-quadruplex structure of the conserved $(GGA)_4$, we prepared several mutants of the $(GGA)_4$ region (m1–m5) and analyzed their affinities compared to bPrP and bPrP-β. The K_d values for different mutants are shown in Table 3. All #1 mutants showed decrease of binding ability to bPrP-β. The conserved $(GGA)_4$ is therefore important for binding to bPrP-β. In the case of interactions with bPrP, all mutants except m4 showed moderate decrease in binding affinity (1.8–2.5-fold). This is due to the ability of PrP to bind to nucleic acids.[29,30]

Since we could not verify the specific binding of mutants to bPrP from the determined K_d values, we carried out competitive binding analysis using the labeled apt #1 in the presence of elevated amounts of unlabeled competitors (Fig. 5). We used non G-quadruplex forming competitor (m5) and probable G-quadruplex forming competitors (m7, m8, apt #1). It was demonstrated that apt #1 was the best competitor, and all probable G-quadruplex forming competitors competed with apt #1 for binding to bPrP in dose-dependent manner. There are small differences of displacement efficiency among apt #1, m7 and m8. By comparing the competing ability of G-quadruplex forming RNAs m7, m8 and apt #1, the adenine residues in $(GGA)_4$ are also important for binding to bPrP. Non G-quadruplex forming competitor (m5) showed weak competition in the presence of K^+ ion. Under the no salt condition m5 competed with apt #1 in the similar level of apt #1 (Fig. 5B). This means that binding ability of apt #1 to bPrP is specific binding in the presence of K^+ ion.

Minimization of apt #1 retained the binding ability to bPrP and bPrP-β. Since the conserved region $(GGA)_4$ of apt #1 plays important role for specific binding to bPrP and bPrP-β, we minimized the aptamer to a 12-nt RNA $r(GGA)_4$ and measured K_d values (Table 4). Compared with apt #1, $r(GGA)_4$ showed better affinity for bPrP (K_d = 8.5 nM) and similar K_d value for bPrP-β (K_d = 280 nM). We then examined $d(GGA)_4$, which forms the unique G-quadruplex structure shown in Figure 3B.[24] This revealed that $d(GGA)_4$ showed a strikingly lower affinity for both bPrP and bPrP-β, compared to $r(GGA)_4$. This result suggests that $r(GGA)_4$ has a different G-quadruplex structure with that of $d(GGA)_4$.

To test the RNA sequence specificity of $r(GGA)_4$, we analyzed the binding of R14, which contains GGA residues and forms another type of intramolecular G-quadruplex consisting of tetrad and hexad in the presence of K^+.[31] As shown in Table 4, R14 showed greater affinity for both proteins compared to $d(GGA)_4$, but lower affinity than $r(GGA)_4$. Hence the conserved tandem GGA repeat $(GGA)_4$ is the important region for binding to bPrP and bPrP-β, and the minimized 12-nt RNA retained this binding ability.

Adenine stretch attached at 5'-site of $(GGA)_4$ enhanced the binding affinity for bPrP-β. Besides of conserved region $(GGA)_4$, adenine tracts (3–6 residues) were frequently observed at the 5'-site of $(GGA)_4$ in obtained RNA aptamers (Table 2). The capacity of these adenine sequences to bind to bPrPs was investigated. The

sequence of constructs and their K_d values are shown in Table 5. Binding affinities were almost the same (K_d = 9–17 nM) for full-length bPrP. For truncated variant bPrP (102–241), they also showed the same level of binding affinities (K_d = 160–270 nM) except r(GGA)$_4$-19 attached 7 adenines (K_d = 540 nM). However with regard to bPrP-β, the five adenine stretch r(GGA)$_4$-17 showed the lowest K_d value (K_d = 78 nM), followed by the four adenines stretch r(GGA)$_4$-16 (K_d = 92 nM). The optimal numbers of adenine correspond with the trend observed in the RNA aptamers.

As reported previously, amino acids 23–90 (human numbering) of PrP interact with nucleic acids in a specific or non-specific manner.[32] Short RNAs including r(GGA)$_4$-17 bound to bPrP (102–241) (probably amino acids residues 102–131 from the binding result of #1 in Fig. 4) in a specific manner. Therefore, like an antibody, r(GGA)$_4$-17 shown in Table 5 may recognize PrP27-30, which is a proteinase K (PK) resistant fragment following removal of the N-terminal region (approximately up to amino acid position 90 in mPrP).[33-36]

Discussion

In this study, we described an in vitro selection of RNA aptamers against recombinant bPrP. All selected RNA aptamers contained the four repeats of a two guanine containing sequence and most of RNAs contained the conserved sequence (GGA)$_4$ (Table 2). Our findings suggest that apt #1 can be used to detect PrP^C in bovine brain homogenate, a process which more conventionally employs an antibody with an immuno-blotting assay (Fig. 2). By collecting CD spectra in increasing concentrations of KCl, apt #1 probably forms a parallel G-quadruplex derived from the (GGA)$_4$ sequence.

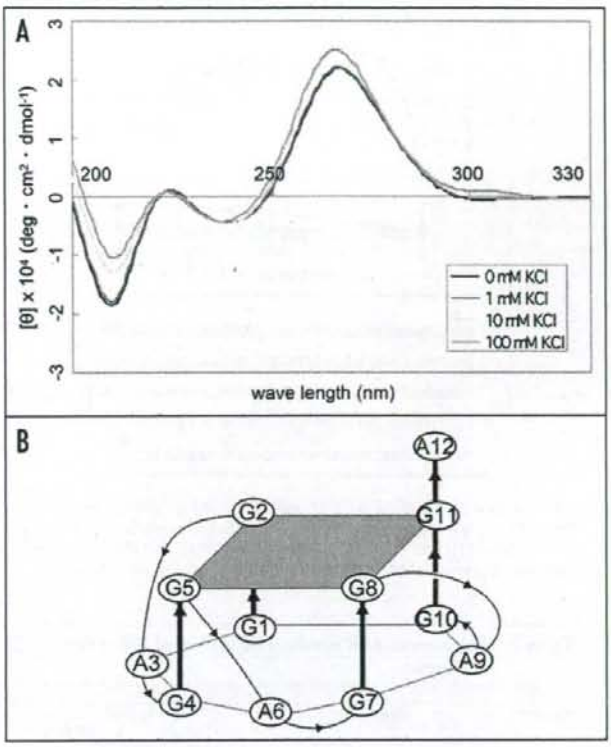

Figure 3. CD spectra of apt #1 and r(GGA)$_4$-15 in the presence of KCl. CD spectra were measured at 20°C in titration with KCl (0, 1, 10, 100 mM). (A) CD spectra of 10 μM apt #1. (B) The structure of d(GGA)$_4$.[24]

It has been reported that the proposed G-quadruplex structure plays a critical role in the binding to PrP.[10,11] Among the reported anti-PrP RNA aptamers containing GGA repeats, anti-syrian golden hamster PrP aptamers contain GGGA repeat interacting with amino acids 23–52,[10] and aptamer DP7 targeted to amino acids 90–129 of human PrP involves GGA repeats.[11] Anti-mPrP aptamer 60–3 used as a competitor in this study also contains GGAGG repeat interacting with amino acids 23–119.[16] The aptamer 60–3 binds to bPrP in the presence of Na^+,[16] however it did not show enhanced bindings to both bPrP and bPrP-β in the presence of K^+ (data not shown). The conserved sequence of tandem GGA repeat (GGA)$_4$ is isolated only in this study. Considering the high frequent appearance of GGA observed in anti-PrP RNA aptamers, GGA repeat must be meaningful for specific binding to PrPs.

We confirmed that the conserved sequence (GGA)$_4$ is important for specific binding to both bPrP and bPrP-β by investigating the effect of substitutions in this region (Table 3) and competitive binding analysis (Fig. 5). Amongst G-quartets which form small DNA and RNAs involving a GGA repeat, r(GGA)$_4$ showed the best affinity to both bPrP and bPrP-β indicating the minimal sequence required for the sequence specificity for preferential binding (Table 4). d(GGA)$_4$ and R14,[31] form similar intramolecular parallel G-quadruplexes consisting of tetrad-heptad and tetrad-hexad, respectively.[24] R14 showed better binding to bPrP and bPrP-β compared to d(GGA)$_4$. This preferential binding may be derived either from the RNA molecule or differences in the tertiary G-quadruplex structures. To clarify the relationships between binding and structure, we are in the process of investigating these structures in more detail.

The binding affinity of bPrP-β with the G_0 RNA pool was notably low (data not shown). The recombinant amyloidogenic prion protein, bPrP-β, is generated through chemical treatment of recombinant bPrP which enhances the formation of the β-sheet secondary structure in the C-terminal region.[37] The N-terminus of PrP (amino acids 23–120 in mPrP) is a flexible non-structural region and the C-terminus of PrP is preserved in the globular three-dimensional structure.[38] It is known that the N-terminal region of PrP interacts with nucleic acids in non-specific and/or specific manner.[29,30,32] In an NMR study of the interaction between murine recombinant PrP and RNA aptamer SAF93,[12,43-59] it was observed that part of the N-terminal domain is modified after interaction with RNA.[39] From these experimental results, it can be deduced that the environment of N-terminal regions also differ between bPrP and bPrP-β. One possibility is that the packed structure of the C-terminal region, which contains higher numbers of β-sheets in bPrP-β, causes a narrowing of free space and a decrease in structural flexibility in the N-terminal region, resulting in poor acceptability of nucleic acids and induced fitting. Finally, we showed that the binding ability for

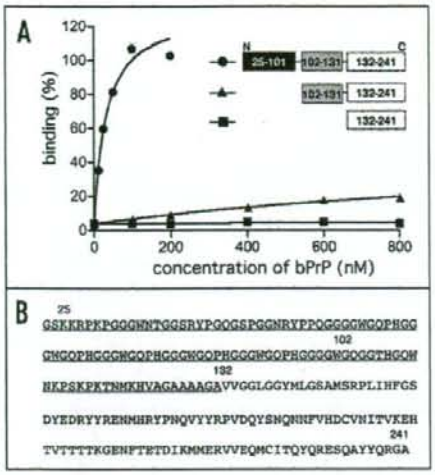

Figure 4. Binding affinity of anti-bPrP aptamer apt #1 for deletion variants of bPrP. (A) Circles, triangles and squares indicate full length bPrP (25–241), bPrP (102–241) and bPrP (132–241), respectively. (B) The amino-acid sequence alignment of bPrP. The apt #1 binding region is underlined.

Table 3 Sequences and binding affinities of different mutants

aptamers	sequence	K_d (nM) bPrP	bPrP-β
apt #1	5'-GGAGGAGGAGGA-3'	31 ± 6	220 ± 67
m1	5'-GGAGGAGGA-3'	79 ± 24	≥1000
m2	5'-GGAGGA-3'	75 ± 28	≥1000
m3	5'-(no GGA repeat)-3'	55 ± 24	≥1000
m4	5'-UUUUUUUUUUU-3'	134 ± 42	≥1000
m5	5'-UUAUUAUUAUUA-3'	56 ± 16	≥1000

bPrP-β could be improved by minimization of apt #1 to a 17 nt RNA containing an adenosine sequence at the 5'-site of $(GGA)_4$.

By investigating the binding properties of deletion variants of bPrP, we identified the N-terminal region of bPrP, amino acids 25–131, as being a key region for apt #1 binding. Apt #1 could bind to amino acids 102–131 of bPrP. It is known that amino acids 25–101 (23–90 in hPrP) contain the nucleic acid binding site.[32] Notably, the additional N-terminal binding region 102–131 provides apt #1 with specific binding activity. Furthermore, it conferred the ability of the minimized aptamer to bind with both bPrP and bPrP-β with high affinity.

Aptamers, with their high ligand specificity, can enrich a target molecule from biological samples. For example, RNA ligands have been successfully used for the concentration of PrP^C and PK-resistant PrP taken from serum, urine[32] and brain homogenate.[40] Previously, we also used anti-mPrP aptamer for bead-based purification and concentration of PrP^C from mouse brain homogenate.[16] Takemura et al. proposed that PrP^{Sc}-enrichment with pre-treatment using PrP^C-specific aptamers, to remove normal prions from a sample, could be applied as diagnostic tools in double ligand assay systems.[17] Such an enrichment method of PrP^{Sc} might be required for early stage detection of the β form PrP because it exists in extremely low proportions compared to the α form. After treatment with PrP^C-specific ligands, apt #1 may be useful for PrP^{Sc} enrichment, rather than employing an antibody without PK treatment.

In a recent publication investigating the interaction between RNAs and recombinant PrP, it was shown that PrP aggregates could be induced by the presence of small synthetic oligonucleotides and aggregates with total RNA extract from cultured mouse neuroblastoma cells (N2a) were cytotoxic.[39] A sensitive method of pathological prion protein detection using cyclic amplification of protein misfolding (PMCA) has been developed.[41] It has been reported that successful PMCA propagation of PrP^{Sc} molecules in a purified system requires accessory poly (A) RNA molecules.[42] Association with PrP and some specific RNAs such as RNA aptamers might help in this field.

In conclusion, we succeeded in harvesting an RNA aptamer against bPrP and bPrP-β, and in producing a high affinity, and a 17 mer minimized form containing 5 adenines at the 5'-site of $(GGA)_4$. Its ability to bind to bPrP (102–241), which corresponds to PrP 27–30, indicates that it may detect PK-digested PrP as well as antibodies. These aptamers will, we believe, be valuable in the study of conversions, diagnosis and therapy of prion disease.

Materials and Methods

Proteins. Recombinant bPrP (amino acids 25–241), amyloidogenic PrP prepared from recombinant bPrP (bPrP-β) and truncated bPrPs [bPrP (102–241): amino acids 102–241; bPrP (132–241): amino acids 132–241] were purchased from Alicon (Switzerland).

In vitro selection of RNA aptamers against bPrP. An N55S RNA pool containing a region of 55 randomized nucleotide [5'-GGGAGGUGGAACUGAAGGAGA-(N55)-ACUUCGCAAUCGCUCUACGCA-3'] was used to perform in vitro selection on a 0.45 μm HAWP nitrocellulose filter (Millipore). RNAs were heat denatured at 90°C for 2 min, 72°C for 5 min, 55°C for 5 min, 37°C for 2 min and cooled to room temperature in 50 μl of binding buffer [20 mM Tris-HCl (pH 7.5) and 100 mM NaCl] prior to use. The RNA pool was incubated for 10 to 30 min with bPrP in the presence of competitor RNAs (tRNA and/or anti-mPrP aptamer 60–3,[16] at room temperature. This mixed solution was passed through the nitrocellulose filter and washed with the binding buffer. RNA bound to bPrP on the filter was recovered with 400 μl of 7 M urea at 90°C for 5 min. The eluted RNA was ethanol precipitated and reverse transcribed using AMV reverse transcriptase (Wako) at 42°C for 1 h. The product was PCR amplified (94°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec) using Gene Taq (Nippon Gene) with forward primer [5'-TGTAATACGACTCACTATAGG GAGGTGGAACTGAAGGAGA-3'] and reverse primer [5'-TGC GTAGAGCGATTGCGAAGT-3'] (Fasmac), and transcribed using the T7 Ampliscribe Kit (Epicentre Technologies). The RNA product was treated with DNase I and purified by Micro Bio-Spin Columns P-30 (Bio-Rad) or 8% PAGE containing 7 M urea and subjected to the next round of selection. From the eighth round of selection mutagenic PCR[21] was introduced. After the tenth generation cDNA pool was inserted into the pGEM-T Easy vector (Promega), cloned in *Escherichia coli* JM109 strain, and sequenced (ABI 3100; Applied Biosystems).

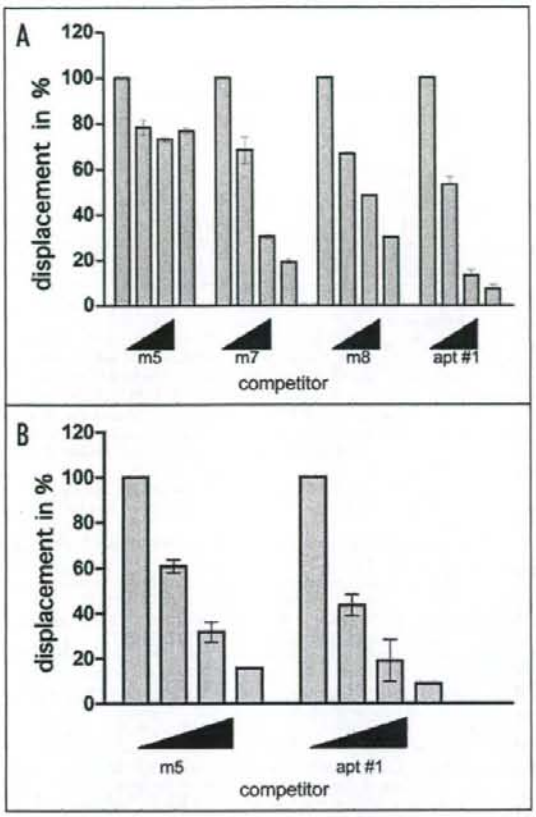

Figure 5. Comparison of binding of apt #1 and bPrP using competitive binding assay. The #1 mutants (m7, m8) contained substitutions (shown underlined) in the conserved region $(GGA)_4$ as follows: m7: 5'-GGUGGAGGAGGA-3'; m8: 5'-GGUGGUGGUGGU-3'. The competitor RNAs were added in various concentrations with 0, 1, 5 and 10-fold greater molar concentrations of labeled apt #1. Binding assay was performed in the presence of 10 mM KCl (A) and in the absence of KCl (B).

Preparation of RNA aptamer. To prepare RNA aptamer, the double-stranded DNA generated by PCR was used as a template for in vitro transcription by T7 RNA polymerase, as described above.

Mutant RNAs of apt #1 were prepared from mutagenic PCR using the following DNA templates (Fasmac) and the selection primers: for m1 [5'-TGC GTA GAG CGA TTG CGA AGT TGC TCA UCC UCC UCC TCT TGG GCT ACT TCC TCA TTC GAG AGA TGA ATG GAT TGT CTC CTT CAG TTC CAC CT CCC-3']; for m2 [5'-TGC GTA GAG CGA TTG CGA AGT TGC TCA UCC UCC TCT TGG GCT ACT TCC TCA TTC GAG AGA TGA ATG GAT TGT CTC CTT CAG TTC CAC CTC CC-3']; for m3 [5'-TGC GTA GAG CGA TTG CGA AGT TGC TCA TCT TGG GCT ACT TCC TCA TTC GAG AGA TGA ATG GAT TGT CTC CTT CAG TTC CAC CTC CC-3']; for m4 [5'-TGC GTA GAG CGA TTG CGA AGT TGC TCA AAA AAA AAA AAA TCT TGG GCT ACT TCC TCA TTC GAG AGA TGA ATG GAT TGT CTC CTT CAG TTC CAC CTC CC-3']; for m5 [5'-TGC GTA GAG CGA TTG CGA AGT TGC TCA TUU TUU TUU TUU TCT TGG GCT ACT TCC TCA TTC GAG AGA TGA ATG GAT TGT CTC CTT CAG TTC CAC CTC CC-3']; for m7 [5'-TGC GTA GAG CGA TTG CGA AGT TGC TCA TCC TCC TCC ACC TCT TGG GCT ACT TCC TCA TTC GAG AGA TGA ATG GAT TGT CTC CTT CAG TTC CAC CTC CC-3']; for m8 [5'-TGC GTA GAG CGA TTG CGA AGT TGC TCA ACC ACC ACC ACC TCT TGG GCT ACT TCC TCA TTC GAG AGA TGA ATG GAT TGT CTC CTT CAG TTC CAC CTC CC-3']. To prepare C#1, the generated PCR fragment using plasmid #1 as template and the proper primers (+) c#1 and (-) c#1, were used for in vitro transcription by T7 RNA polymerase as described above. (+) c#1: [5'-TGT AAT ACG ACT CAC TAT AGG CGT AGA GCG ATT GCG AAG-3']; and (-) c#1: [5'-GGG AGG TGG AAC TGA AGG AGA-3'].

The chemically synthesized RNAs or DNAs: $r(GGA)_4$-15, $r(GGA)_4$, $d(GGA)_4$, R14, $r(GGA)_4$-16, $r(GGA)_4$-17, $r(GGA)_4$-18 and $r(GGA)_4$-19 were purchased from Fasmac (Japan). 5'-biotinylated RNAs (Bi#1, BiC#1) were prepared with a 5' end tag nucleic acid labeling system (Vector Laboratories).

Binding assay of anti-bPrP aptamer. Radioisotope labeling of RNA by in vitro transcription was carried out using α-^{32}P-ATP, as previously described.[16] Refolded ^{32}P labeled aptamer (10 nM) was mixed with varying concentrations of bPrP, or its derivatives, to a total volume of 25 µl in reaction buffer [20 mM Tris-HCl (pH 7.5), 100 mM NaCl or 10 mM KCl]. After 20 min incubation, the mixture was passed through a nitrocellulose filter and washed with 500 µl of the reaction buffer. The amount of bound RNA was measured with BAS 2500 (Fuji Film), and binding activities were calculated as the percentage of input RNA retained on the filter in the protein-RNA complex. We determined the equilibrium dissociation constant (K_d) using GraphPad PRISM using non-linear regression curve fitting, and a one site binding hyperbola equation (RNA binding (%) = B_{max} x [PrP]/(K_d + [PrP]), where B_{max} is the maximum bound at saturating PrP concentrations).

Northwestern assay with the aptamer. The Northwestern assay was performed using conventional methods.[22] A sample of 10 µl 10% bovine brain homogenate in suspension buffer [0.1% Nonidet P-40, 0.1% deoxycholate, 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl] was separated by 10% SDS-PAGE and transferred to a 0.22 µm nitrocellulose membrane (Bio-Rad). A sample of 2 ml 1% BSA in binding buffer [20 mM Tris-HCl (pH 7.5) and 10 mM KCl] was used for blocking for 30 min. Proceeding this were 30 min incubation, with 200 nM of 5'-biotin labeled apt #1

Table 4 **Comparison of binding affinities for bPrP and bPrP-β between GGA repeat containing RNAs and DNA**

RNA/DNA	sequence	K_d (nM)	
		bPrP	bPrP-β
apt #1	-	31 ± 6	220 ± 70
$r(GGA)_4$	r(GGAGGAGGAGGA)	8.5 ± 3.4	280 ± 80
$d(GGA)_4$	d(GGAGGAGGAGGA)	64 ± 31	≥1000
R14	r(GGAGGUUUUGGAGG)	32 ± 18	610 ± 240

Table 5 **Comparison of adenine stretch at 5'-site of $(GGA)_4$ with binding affinities of bPrP, bPrP(102-241) and bPrP-β**

aptamers	sequence	K_d (nM) bPrP	bPrP (102-241)	bPrP-β
r$(GGA)_4$-19	AAAAAAAGGAGGAGGAGGA	17 ± 5	540 ± 310	270 ± 90
r$(GGA)_4$-18	AAAAAAGGAGGAGGAGGA	10 ± 2	160 ± 50	150 ± 40
r$(GGA)_4$-17	AAAAAGGAGGAGGAGGA	9.4 ± 4.0	270 ± 120	78 ± 30
r$(GGA)_4$-16	AAAAGGAGGAGGAGGA	16 ± 3	190 ± 50	92 ± 26
r$(GGA)_4$-15	AAA GGAGGAGGAGGA	16 ± 4	220 ± 90	200 ± 90
r$(GGA)_4$	GGAGGAGGAGGA	8.5 ± 3.4	240 ± 90	280 ± 80

(Bi#1) in binding buffer, followed by a 30 min incubation with streptavidin-alkaline phosphatase conjugate (SA-AP; 2 ng/μl; Roche Applied Science) in binding buffer and a 1 min incubation with CDP-Star (20 mg/ml, Roche Applied Science). Detection was carried out using ECL Mini-Camera (Amersham Biosciences) and Polaroid film. For the standard immuno-blotting assay of $bPrP^C$, a T2 antibody was used.[23] All procedures were performed at room temperature.

Dot-blotting assay. 5' biotinylated RNA, Bi#1 or BiC#1, was spotted on nitrocellulose membrane and UV cross-linked for 5 min. Different amounts of bPrP (480, 240, 120, 60 and 30 ng) were spotted and air-dried for 10 min. The membrane was treated by blocking with 1 % BSA in reaction buffer [20 mM Tris-HCl (pH 7.5), 10 mM KCl] for 30 min, incubating with 200 nM of Bi#1 or BiC#1 in the presence of poly U (0.5 μg/μl, Amersham Biosciences) for 20 min, and then incubating with SA-AP (2 ng/μl) in reaction buffer for 30 min. Following a 1 min incubation with CDP-Star (20 mg/ml) in 0.1 M Tris-HCl (pH 9.5), detection was carried out with ECL Mini-Camera and Polaroid film. All procedures were performed at room temperature.

Circular dichroism spectroscopy. CD spectra were recorded with a JASCO J-720 Spectropolarimeter (JASCO) as previously reported.[24] A cell of 1 mm light path length and 300 μl volume was used to analyze in titration with KCl (0, 1, 10, 100 mM). The spectra were scanned four times from 200 nm to 330 nm. The CD intensities were expressed in [θ] per residue.

Competitive binding assay against bPrP. To characterize the specific binding of apt #1, binding of apt #1 to bPrP with increasing concentrations of competitor RNAs or DNAs were measured using a filter-binding assay. The labeled apt #1 (25 nM) was incubated with bPrP (50 nM) in the presence of a competitor (0, 1, 5 and 10-fold molar ratios) in binding buffer [20 mM Tris-HCl (pH 7.5), 10 mM KCl] for 20 min. The mixture was passed through a nitrocellulose filter, washed with 500 μl binding buffer, and the amount of binding of labeled apt #1 was determined as described above. 100% maximal binding is defined as the percentage of apt #1 bound to bPrP in the absence of any competitor.

Acknowledgements

The authors thank Drs. Katahira and Matsugami of Yokohama City University for measurement of CD spectra and helpful discussions. This work was supported by the fund from National Institute of Advanced Industrial Science and Technology (to S.N.) and by Grant-in-Aid from the BSE Control Project of the Ministry of Agriculture, Forestry and Fisheries of Japan (to T.Y.).

References

1. Pan KM, Baldwin M, Nguyen J, Gasset M, Serban A, Groth D, Mehlhorn I, Huang Z, Fletterick RJ, Cohen FE, et al. Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. Proc Natl Acad Sci USA 1993; 90:10962-6.
2. Prusiner SB. Prions. Proc Natl Acad Sci USA 1998; 95:13363-83.
3. Marc D, Mercey R, Lantier F. Scavenger, transducer, RNA chaperone? What ligands of the prion protein teach us about its function. Cell Mol Life Sci 2007; 64:815-29.
4. Wadsworth JD, Collinge J. Update on human prion disease. Biochim Biophys Acta 2007; 1772:598-609.
5. Hu W, Kieseier B, Frohman E, Eagar TN, Rosenberg RN, Hartung HP, Stuve O. Prion proteins: physiological functions and role in neurological disorders. J Neurol Sci 2008; 264:1-8.
6. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. Nature 1990; 346:818-22.
7. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 1990; 249:505-10.
8. Yang Y, Yang D, Schluesener HJ, Zhang Z. Advances in SELEX and application of aptamers in the central nervous system. Biomol Eng 2007; 24:583-92.
9. Mairal T, Ozalp VC, Lozano Sanchez P, Mir M, Katakis I, O'Sullivan CK. Aptamers: molecular tools for analytical applications. Anal Bioanal Chem 2008; 390:989-1007.
10. Weiss S, Proske D, Neumann M, Groschup MH, Kretzschmar HA, Famulok M, Winnacker EL. RNA aptamers specifically interact with the prion protein PrP. J Virol 1997; 71:8790-7.
11. Proske D, Gilch S, Wopfner F, Schatzl HM, Winnacker EL, Famulok M. Prion-protein-specific aptamer reduces PrP^{Sc} formation. Chembiochem 2002; 3:717-25.
12. Rhie A, Kirby L, Sayer N, Wellesley R, Disterer P, Sylvester I, Gill A, Hope J, James W, Tahiri-Alaoui A. Characterization of 2'-fluoro-RNA aptamers that bind preferentially to disease-associated conformations of prion protein and inhibit conversion. J Biol Chem 2003; 278:39697-705.
13. Sayer NM, Cubin M, Rhie A, Bullock M, Tahiri-Alaoui A, James W. Structural determinants of conformationally selective, prion-binding aptamers. J Biol Chem 2004; 279:13102-9.
14. Sekiya S, Nishikawa F, Noda K, Kumar PK, Yokoyama T, Nishikawa S. In vitro selection of RNA aptamers against cellular and abnormal isoform of mouse prion protein. Nucleic Acids Symp Ser (Oxf) 2005; 361-2.
15. Mercey R, Lantier I, Maurel MC, Grosclaude J, Lantier F, Marc D. Fast, reversible interaction of prion protein with RNA aptamers containing specific sequence patterns. Arch Virol 2006; 151:2197-214.
16. Sekiya S, Noda K, Nishikawa F, Yokoyama T, Kumar PK, Nishikawa S. Characterization and application of a novel RNA aptamer against the mouse prion protein. J Biochem 2006; 139:383-90.
17. Takemura K, Wang P, Vorberg I, Surewicz W, Priola SA, Kanthasamy A, Pottathil R, Chen SG. Sreevatsan S. DNA aptamers that bind to PrP(C) and not PrP(Sc) show sequence and structure specificity. Exp Biol Med (Maywood) 2006; 231:204-14.
18. Nishikawa F, Murakami K, Noda K, Yokoyama T, Nishikawa S. Detection of structural changes of RNA aptamer containing GGA repeats under the ionic condition using the microchip electrophoresis. Nucleic Acids Symp Ser (Oxf) 2007; 397-8.
19. Bibby DF, Gill AC, Kirby L, Farquhar CF, Bruce ME, Garson JA. Application of a novel in vitro selection technique to isolate and characterise high affinity DNA aptamers binding mammalian prion proteins. J Virol Methods 2008; 151:107-15.
20. Ogasawara D, Hasegawa H, Kaneko K, Sode K, Ikebukuro K. Screening of DNA Aptamer Against Mouse Prion Protein by Competitive Selection. Prion 2008; 1:248-54.
21. Fromant M, Blanquet S, Plateau P. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. Anal Biochem 1995; 224:347-53.
22. Bagga PS, Wilusz J. Northwestern screening of expression libraries. Methods Mol Biol 1999; 118:245-56.
23. Hayashi H, Takata M, Iwamaru Y, Ushiki Y, Kimura KM, Tagawa Y, Shinagawa M, Yokoyama T. Effect of tissue deterioration on postmortem BSE diagnosis by immunobiochemical detection of an abnormal isoform of prion protein. J Vet Med Sci 2004; 66:515-20.
24. Matsugami A, Ouhashi K, Kanagawa M, Liu H, Kanagawa S, Uesugi S, Katahira M. An intramolecular quadruplex of (GGA)(4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction. J Mol Biol 2001; 313:255-69.

25. Laughlan G, Murchie AI, Norman DG, Moore MH, Moody PC, Lilley DM, Luisi B. The high-resolution crystal structure of a parallel-stranded guanine tetraplex. Science 1994; 265:520-4.
26. Davis JT. G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. Angew Chem Int Ed Engl 2004; 43:668-98.
27. Balagurumoorthy P, Brahmachari SK, Mohanty D, Bansal M, Sasisekharan V. Hairpin and parallel quartet structures for telomeric sequences. Nucleic Acids Res 1992; 20:4061-7.
28. Catasti P, Chen X, Moyzis RK, Bradbury EM, Gupta G. Structure-function correlations of the insulin-linked polymorphic region. J Mol Biol 1996; 264:534-45.
29. Nandi PK. Interaction of prion peptide HuPrP106-126 with nucleic acid. Arch Virol 1997; 142:2537-45.
30. Nandi PK, Nicole JC. Nucleic acid and prion protein interaction produces spherical amyloids which can function in vivo as coats of spongiform encephalopathy agent. J Mol Biol 2004; 344:827-37.
31. Liu H, Kugimiya A, Matsugami A, Katahira M, Uesugi S. Quadruplex structures of RNA 14-mer, r(GGAGGUUUUGGAGG) and DNA 14-mer, d(GGAGGTTTTGGAGG). Nucleic Acids Res Suppl 2002; 2:177-8.
32. Zeiler B, Adler V, Kryukov V, Grossman A. Concentration and removal of prion proteins from biological solutions. Biotechnol Appl Biochem 2003; 37:173-82.
33. Prusiner SB, McKinley MP, Bowman KA, Bolton DC, Bendheim PE, Groth DF, Glenner GG. Scrapie prions aggregate to form amyloid-like birefringent rods. Cell 1983; 35:349-58.
34. Prusiner SB, Groth DF, Bolton DC, Kent SB, Hood LE. Purification and structural studies of a major scrapie prion protein. Cell 1984; 38:127-34.
35. Caughey B, Raymond GJ, Ernst D, Race RE. N-terminal truncation of the scrapie-associated form of PrP by lysosomal protease(s): implications regarding the site of conversion of PrP to the protease-resistant state. J Virol 1991; 65:6597-603.
36. Prusiner SB, McKinley MP, Groth DF, Bowman KA, Mock NI, Cochran SP, Masiarz FR. Scrapie agent contains a hydrophobic protein. Proc Natl Acad Sci USA 1981; 78:6675-9.
37. Luhrs T, Zahn R, Wuthrich K. Amyloid formation by recombinant full-length prion proteins in phospholipid bicelle solutions. J Mol Biol 2006; 357:833-41.
38. Riek R, Hornemann S, Wider G, Glockshuber R, Wuthrich K. NMR characterization of the full-length recombinant murine prion protein, mPrP(23-231). FEBS Lett 1997; 413:282-8.
39. Gomes MP, Millen TA, Ferreira PS, Cunha ESNL, Vieira TC, Almeida MS, Silva JL, Cordeiro Y. Prion protein complexed to N2a cellular RNAs through its N-terminal domain forms aggregates and is toxic to murine neuroblastoma cells. J Biol Chem 2008; 283:19616-25.
40. Davidowitz E, Eljuga L, Dover K, Tian J, Grossman A. Concentration of prion protein from biological samples to increase the limits of detection by immunoassay. Biotechnol Appl Biochem 2005; 41:247-53.
41. Saborio GP, Permanne B, Soto C. Sensitive detection of pathological prion protein by cyclic amplification of protein misfolding. Nature 2001; 411:810-3.
42. Deleault NR, Harris BT, Rees JR, Supattapone S. Formation of native prions from minimal components in vitro. Proc Natl Acad Sci USA 2007; 104:9741-6.