In summary, we established an in vitro mouse NK1/NK2 cell differentiation system and revealed possible molecular events that may control the chromatin remodeling and transcriptional activation of the type 2 cytokine gene loci in NK2 cells. Low level expression of GATA3 and high level expression of ROG may confer unique cytokine production profiles on NK2 cells. It is now of interest to study the physiological roles of NK2 cells possessing unique type 2 cytokine production profiles in murine models of infection and various immune disorders in vivo.

## Acknowledgments

We thank Kaoru Sugaya for excellent technical assistance.

## References


1. Miller, J. S. 2001. The biology of natural killer cells in cancer, infection, and pregnancy. *Exp. Hematol. 29:1157.*
2. Smyth, M. J., Y. Hayakawa, K. Takeda, and H. Yagita. 2002. New aspects of natural-killer-cell surveillance and therapy of cancer. *Nat. Rev. Cancer 2:850.*
3. Warren, H. S., B. F. Kinnear, J. H. Phillips, and L. L. Lanier. 1995. Production of IL-5 by human NK cells and regulation of IL-5 secretion by IL-4, IL-10, and IL-12. *J. Immunol. 154:5144.*
4. Peritt, D., S. Robertson, G. Gri, L. Showe, M. Aste-Amezaga, and G. Trinchieri. 1998. Differentiation of human NK cells into NK1 and NK2 subsets. *J. Immunol. 161:5821.*
5. Loza, M. J., L. Zamai, L. Azzoni, E. Rosati, and B. Perussia. 2002. Expression of type 1 (interferon γ) and type 2 (interleukin-13, interleukin-5) cytokines at distinct stages of natural killer cell differentiation from progenitor cells. *Blood 99:1273.*
6. Deniz, G., M. Akdis, E. Aktas, K. Blaser, and C. A. Akdis. 2002. Human NK1 and NK2 subsets determined by purification of IFN-γ-secreting and IFN-γ-non-secreting NK cells. *Eur. J. Immunol. 32:879.*
7. Hoshino, T., R. T. Winkler-Pickett, A. T. Mason, J. R. Ortaldo, and H. A. Young. 1999. IL-13 production by NK cells: IL-13-producing NK and T cells are present in vivo in the absence of IFN-γ. *J. Immunol. 162:51.*
8. Walker, C., J. Checkel, S. Cammisuli, P. J. Leibson, and G. J. Gleich. 1998. IL-5 production by NK cells contributes to eosinophil infiltration in a mouse model of allergic inflammation. *J. Immunol. 161:1962.*
9. Korsgren, M., C. G. Persson, F. Sundler, T. Bjerke, T. Hansson, B. J. Chambers, S. Hong, L. Van Kaer, H. G. Ljunggren, and O. Korsgren. 1999. Natural killer cells determine development of allergen-induced eosinophilic airway inflammation in mice. *J. Exp. Med. 189:553.*
10. Takahashi, K., S. Miyake, T. Kondo, K. Terao, M. Hatakenaka, S. Hashimoto, and T. Yamamura. 2001. Natural killer type 2 bias in remission of multiple sclerosis. *J. Clin. Invest. 107:R23.*
11. Abbas, A. K., K. M. Murphy, and A. Sher. 1996. Functional diversity of helper T lymphocytes. *Nature 383:787.*
12. Nelms, K., A. D. Keegan, J. Zamorano, J. J. Ryan, and W. E. Paul. 1999. The IL-4 receptor: signaling mechanisms and biologic functions. *Annu. Rev. Immunol. 17:701.*
13. Murphy, K. M., W. Ouyang, J. D. Farrar, J. Yang, S. Ranganath, H. Asnagli, M. Afkarian, and T. L. Murphy. 2000. Signaling and transcription in T helper development. *Annu. Rev. Immunol. 18:451.*
14. O'Garra, A. 1998. Cytokines induce the development of functionally heterogeneous T helper cell subsets. *Immunity 8:275.*
15. Constant, S. L., and K. Bottomly. 1997. Induction of Th1 and Th2 CD4$^+$ T cell responses: the alternative approaches. *Annu. Rev. Immunol. 15:297.*
16. Yamashita, M., M. Kimura, M. Kubo, C. Shimizu, T. Tada, R. M. Perlmutter, and T. Nakayama. 1999. T cell antigen receptor-mediated activation of the Ras/mitogen-activated protein kinase pathway controls interleukin 4 receptor function and type-2 helper T cell differentiation. *Proc. Natl. Acad. Sci. USA 96:1024.*
17. Yamashita, M., M. Katsumata, M. Iwashima, M. Kimura, C. Shimizu, T. Kamata, T. Shin, N. Seki, S. Suzuki, M. Taniguchi, et al. 2000. T cell receptor-induced calcineurin activation regulates T helper type 2 cell development by modifying the interleukin 4 receptor signaling complex. *J. Exp. Med. 191:1869.*
18. Zhang, D. H., L. Cohn, P. Ray, K. Bottomly, and A. Ray. 1997. Transcription factor GATA-3 is differentially expressed in murine Th1 and Th2 cells and controls Th2-specific expression of the interleukin-5 gene. *J. Biol. Chem. 272:21597.*
19. Zheng, W., and R. A. Flavell. 1997. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell 89:587.*
20. Ouyang, W., S. H. Ranganath, K. Weindel, D. Bhattacharya, T. L. Murphy, W. C. Sha, and K. M. Murphy. 1998. Inhibition of Th1 development mediated by GATA-3 through an IL-4-independent mechanism. *Immunity 9:745.*
21. Szabo, S. J., S. T. Kim, G. L. Costa, X. Zhang, C. G. Fathman, and L. H. Glimcher. 2000. A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell 100:655.*
22. Agarwal, S., and A. Rao. 1998. Long-range transcriptional regulation of cytokine gene expression. *Curr. Opin. Immunol. 10:345.*
23. Strahl, B. D., and C. D. Allis. 2000. The language of covalent histone modifications. *Nature 403:41.*
24. Yamashita, M., M. Ukai-Tadenuma, M. Kimura, M. Omori, M. Inami, M. Taniguchi, and T. Nakayama. 2002. Identification of a conserved GATA3 response element upstream proximal from the interleukin-13 gene locus. *J. Biol. Chem. 277:42399.*
25. Avni, O., D. Lee, F. Macian, S. J. Szabo, L. H. Glimcher, and A. Rao. 2002. $T_H$ cell differentiation is accompanied by dynamic changes in histone acetylation of cytokine genes. *Nat. Immunol. 3:643.*
26. Fields, P. E., S. T. Kim, and R. A. Flavell. 2002. Cutting edge: changes in histone acetylation at the *IL-4* and *IFN-γ* loci accompany Th1/Th2 differentiation. *J. Immunol. 169:647.*
27. Takeda, K., T. Tanaka, W. Shi, M. Matsumoto, M. Minami, S. Kashiwamura, K. Nakanishi, N. Yoshida, T. Kishimoto, and S. Akira. 1996. Essential role of Stat6 in IL-4 signalling. *Nature 380:627.*
28. Nakayama, T., C. H. June, T. I. Munitz, M. Sheard, S. A. McCarthy, S. O. Sharrow, L. E. Samelson, and A. Singer. 1990. Inhibition of T cell receptor expression and function in immature CD4$^+$CD8$^+$ cells by CD4. *Science 249:1558.*
29. Kawano, T., J. Cui, Y. Koezuka, I. Toura, Y. Kaneko, K. Motoki, H. Ueno, R. Nakagawa, H. Sato, E. Kondo, et al. 1997. CD1d-restricted and TCR-mediated activation of Vα14 NKT cells by glycosylceramides. *Science 278:1626.*
30. Omori, M., M. Yamashita, M. Inami, M. Ukai-Tadenuma, M. Kimura, Y. Nigo, H. Hosokawa, A. Hasegawa, M. Taniguchi, and T. Nakayama. 2003. CD8 T cell-specific downregulation of histone hyperacetylation and gene activation of the IL-4 gene locus by ROG, repressor of GATA. *Immunity 19:281.*
31. Kimura, M., Y. Koseki, M. Yamashita, N. Watanabe, C. Shimizu, T. Katsumoto, T. Kitamura, M. Taniguchi, H. Koseki, and T. Nakayama. 2001. Regulation of Th2 cell differentiation by *mel-18*, a mammalian *polycomb* group gene. *Immunity 15:275.*
32. Hattori, M., T. Sudo, H. Izawa, S. Kano, and N. Minato. 1989. Developmental regulation of the extrathymic differentiation potential of the progenitor cells for T cell lineage. *Int. Immunol. 1:151.*
33. Aramburu, J., L. Azzoni, A. Rao, and B. Perussia. 1995. Activation and expression of the nuclear factors of activated T cells, NFATp and NFATc, in human natural killer cells: regulation upon CD16 ligand binding. *J. Exp. Med. 182:801.*
34. Trotta, R., K. A. Puorro, M. Paroli, L. Azzoni, B. Abebe, L. C. Eisenlohr, and B. Perussia. 1998. Dependence of both spontaneous and antibody-dependent, granule exocytosis-mediated NK cell cytotoxicity on extracellular signal-regulated kinases. *J. Immunol. 161:6648.*
35. Kishikawa, H., J. Sun, A. Choi, S. C. Miaw, and I. C. Ho. 2001. The cell type-specific expression of the murine *IL-13* gene is regulated by GATA-3. *J. Immunol. 167:4414.*
36. Lavenu-Bombled, C., C. D. Trainor, I. Makeh, P. H. Romeo, and I. Max-Audit. 2002. Interleukin-13 gene expression is regulated by GATA-3 in T cells: role of a critical association of a GATA and two GATG motifs. *J. Biol. Chem. 277:18313.*
37. Miaw, S. C., A. Choi, E. Yu, H. Kishikawa, and I. C. Ho. 2000. ROG, repressor of GATA, regulates the expression of cytokine genes. *Immunity 12:323.*
38. Inami, M., M. Yamashita, Y. Tenda, A. Hasegawa, M. Kimura, K. Hashimoto, N. Seki, M. Taniguchi, and T. Nakayama. 2004. CD28 costimulation controls histone hyperacetylation of the IL-5 gene locus in developing Th2 cells. *J. Biol. Chem. 279:23123.*
39. Mehrotra, P. T., R. P. Donnelly, S. Wong, H. Kanegane, A. Geremew, H. S. Mostowski, K. Furuke, J. P. Siegel, and E. T. Bloom. 1998. Production of IL-10 by human natural killer cells stimulated with IL-2 and/or IL-12. *J. Immunol. 160:2637.*
40. Yu, C. R., R. A. Kirken, M. G. Malabarba, H. A. Young, and J. R. Ortaldo. 1998. Differential regulation of the Janus kinase-STAT pathway and biologic function of IL-13 in primary human NK and T cells: a comparative study with IL-4. *J. Immunol. 161:218.*
41. Perussia, B., and M. J. Loza. 2003. Linear "2-0-1" lymphocyte development: hypotheses on cellular bases for immunity. *Trends Immunol. 24:235.*
42. Loza, M. J., and B. Perussia. 2001. Final steps of natural killer cell maturation: a model for type 1-type 2 differentiation? *Nat. Immunol. 2:917.*

Biol. Chem., Vol. 384, pp. 1173–1183, August 2003 

# Determination of the NMR Structure of Gln25-Ribonuclease $T_1$

**Ken-ichi Hatano[1,a,*], Masaki Kojima[2], Ei-ichiro Suzuki[3], Masaru Tanokura[4] and Kenji Takahashi[2]**

[1] Department of Biological Sciences, Faculty of Engineering, Gunma University, Kiryu, Gunma 376-8515, Japan
[2] Department of Molecular Biochemistry, School of Life Science, Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo 192-0392, Japan
[3] Central Research Laboratories, Ajinomoto Co., Inc., Suzuki-cho, Kawasaki 210, Japan
[4] Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan

*Corresponding author

**Ribonuclease (RNase) $T_1$ is a guanyloribonuclease, having two isozymes in nature, Gln25- and Lys25-RNase $T_1$. Between these two isozymes, there is no difference in catalytic activity and three-dimensional structure; however, Lys25-RNase $T_1$ is slightly more stable than Gln25-RNase $T_1$. Recently, it has been suggested that the existence of a salt bridge between Lys25 and Asp29/Glu31 in Lys25-RNase $T_1$ contributes to the stability. To elucidate the effects of the replacement of Lys25 with a Gln on the conformation and microenvironments of RNase $T_1$ in detail, the three-dimensional solution structure of Gln25-RNase $T_1$ was determined by simulated-annealing calculations. As a result, the topology of the overall folding was shown to be very similar to that of the Lys25-isozyme except for some differences. In particular, there were two differences in the property of torsion angles of the two disulfide bonds and the conformations of the residues 11–13, 63–66, and 92–93. With regard to the residues 11–13, the lack of the above-mentioned salt bridge in Gln25-RNase $T_1$ was thought to induce the conformational difference of this segment as compared with the Lys25-isozyme. Furthermore, it was proposed that the perturbation of this segment might transfer to the residues 92–93 *via* the two disulfide bonds.**

*Key words:* Electrostatic interaction/Guanyloribonuclease/Nuclear magnetic resonance/$pK_a$ calculation/Ribonuclease $T_1$/Structure calculation.

[a] Present address: Department of Chemistry, Iowa State University, Gilman Hall 0108, Ames, IA 50011, USA

## Introduction

Ribonuclease (RNase) $T_1$ (EC 3.1.27.3) is an acidic protein of 104 amino acid residues, secreted by the fungus *Aspergillus oryzae*. It specifically cleaves a single-stranded RNA chain at the 3'-phosphates of guanylyl residues through transphosphorylation of the 3', 5'-phosphate bonds to form the 2', 3'-cyclic phosphates, followed by hydrolysis of the latter to the corresponding 3'-phosphates (Egami *et al.*, 1964; Takahashi and Moore, 1982; Kojima *et al.*, 2000; Loverix and Steyaert, 2001; Yoshida, 2001). Since RNase $T_1$ is one of the smallest enzymes with a clear base specificity as well as high stability, it has been extensively studied by various methods including chemical modifications (Takahashi and Moore, 1982), protein engineering (Nishikawa *et al.*, 1987; McNutt *et al.*, 1990; Steyaert *et al.*, 1990; Steyaert *et al.*, 1991; Steyaert and Wyns, 1993; Landt *et al.*, 1997; Steyaert, 1997; De Vos *et al.*, 1998), nuclear magnetic resonance (NMR) measurements (Arata *et al.*, 1979; Inagaki *et al.*, 1981; Kyogoku *et al.*, 1982; Hoffmann and Rüterjans, 1988; Shimada and Inagaki, 1990; Schmidt *et al.*, 1991; Fushman *et al.*, 1994; Weisemann *et al.*, 1994; Pfeiffer *et al.*, 1997), and X-ray crystallography (Heinemann and Saenger, 1982; Arni *et al.*, 1988, 1992; Sugio *et al.*, 1988; Hakoshima *et al.*, 1992; Gohda *et al.*, 1994) to elucidate the catalytic mechanism and the mode of interaction with substrate analogs.

Two isoforms of RNase $T_1$ exist in nature: one has Gln and the other Lys at position 25 (Egami *et al.*, 1964; Heinemann and Saenger, 1982). Of these, Gln25-RNase $T_1$ has been the major target of most biochemical and biophysical studies (Takahashi and Moore, 1982), although the three-dimensional (3-D) structure was first determined for Lys25-RNase $T_1$ by X-ray crystallography (Heinemann and Saenger, 1982) and NMR spectroscopy (Pfeiffer *et al.*, 1997). No difference has been detected in catalytic activity between these enzymes; however, the folded structure of the Lys25-isoform was reported to be more stable (by 0.9 kcal/mol) than that of the Gln25-isozyme (Shirley *et al.*, 1989; Kiefhaber *et al.*, 1990, Yu *et al.*, 1994). The conformation of Gln25-RNase $T_1$ determined by X-ray crystallography is indistinguishable from that of Lys25-RNase $T_1$ at a resolution of approx. 1.8 Å, but the location of the base-binding site is fairly different from molecule to molecule in a crystal (Hakoshima *et al.*, 1992). The enzyme contains one α-helix, two isolated β-strands located in the N-terminal region, and a five-stranded anti-parallel β-sheet, and the base recognition site includes Tyr42 through Glu46 and Asn98 (Heinemann and Saenger, 1982; Arni *et al.*, 1988, 1992; Sugio *et al.*, 1988; Hakoshima *et al.*, 1992; Gohda *et al.*, 1994).

Recently, high-resolution $^1$H NMR and multi-dimensional NMR techniques have been successfully applied to the determination of the higher-order structures of proteins. As for RNase $T_1$, a series of NMR studies were performed with the Gln25-isozyme, including investigation of the enzyme-inhibitor interaction (Inagaki *et al.*, 1981; Kyogoku *et al.*, 1982; Hoffmann and Rüterjans, 1988). However, the determination of the solution structure has so far been carried out only for the Lys25-isozyme (Pfeiffer *et al.*, 1997). In the previous study, our group reported the sequence-specific resonance assignments of the $^1$H NMR spectra of Gln25-RNase $T_1$ (Kojima *et al.*, 1995). Comparison of the proton chemical shifts of Gln25-RNase $T_1$ with those of the Lys25-isozyme indicated that the replacement of Lys25 with Gln significantly affects the chemical shifts of the C-terminal part of the α-helix region including Lys/Gln25 and the guanine-binding site (Kojima *et al.*, 1995). Therefore, it is of interest to investigate these effects of the Lys25-to-Gln replacement on the microenvironments of the 3-D solution structure.

In this study, the 3-D solution structure of Gln25-RNase $T_1$ was determined by two-dimensional (2-D) proton NMR spectroscopy at 600 MHz on the basis of 1340 interproton distance constraints derived from the nuclear Overhauser effect (NOE), 30 distance constraints for 15 hydrogen bonds and 37 dihedral angle constraints from coupling constants, and the results were compared with those of the Lys25-isozyme. In addition, we calculated the theoretical $pK_a$ ($pK_{calc}$) values using the atomic coordinates of the structure determined and compared the $pK_{calc}$ values of the catalytically important residues with the experimentally determined $pK_a$ ($pK_{exp}$) values (Inagaki *et al.*, 1981; Spitzner *et al.*, 2001).

## Results

### Structure Calculations of Gln25-RNase $T_1$

A total of 1340 NOE distance constraints were derived from assigned NOE cross peaks, which included 392 intra-residue, 384 sequential ($|i-j| = 1$), 180 medium range ($1 < |i-j| < 5$) and 384 long-range ($|i-j| \geq 5$) constraints (Figure 1). A total of 50 structures were calculated on the

**Table 1** Structural Statistics for the Calculated Structures of Gln25-RNase $T_1$.

| Parameter | Value |
|---|---|
| X-PLOR energies (kcal mol$^{-1}$) | |
| $E_{TOTAL}$ | 243.3±10.2 |
| $E_{BOND}$ | 11.6±0.9 |
| $E_{ANGLE}$ | 145.3±5.3 |
| $E_{IMPR}$ | 19.8±1.2 |
| $E_{VDW}$ | 21.6±3.2 |
| $E_{NOE}$ | 43.3±4.4 |
| $E_{CDIH}$ | 1.7±0.5 |
| RMS deviation from idealized geometry and experimental constraints | |
| Bonds (Å) | 0.0028±0.0001 |
| Angles (deg) | 0.6030±0.0110 |
| Improper torsions (deg) | 0.3950±0.0140 |
| Interproton distances (Å) | 0.0250±0.0010 |
| Torsion angles (deg) | 0.8630±0.1330 |
| Cartesian coordinate RMS differences for backbone atoms (Å) | |
| All backbone atoms (1–104) | 0.85±0.14 |
| Well-defined regions (2–69, 74–103) | 0.77±0.13 |
| α-helix (13–29) | 0.27±0.06 |
| All heavy atoms (1–104) | 1.21±0.11 |

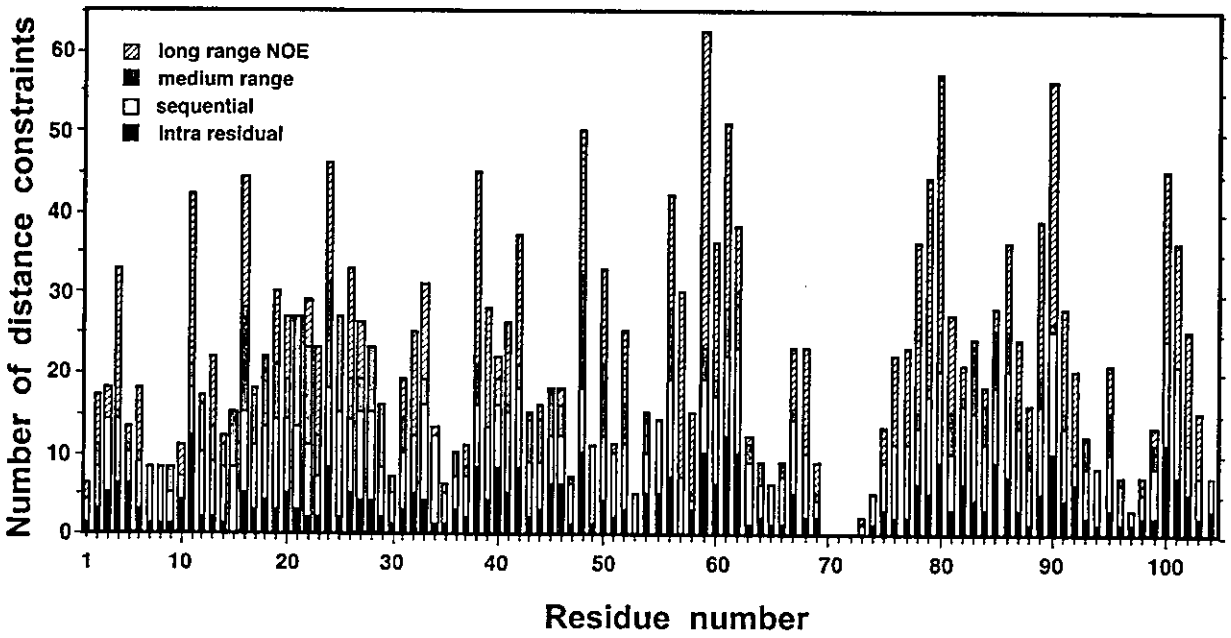


**Fig. 1** Distribution of the Number of Experimental Distance Constraints on the Primary Structure of Gln25-RNase $T_1$.
Filled bars: intra-residual NOEs; open bars: sequential NOEs; shaded bars: medium-range NOEs; hatched bars: long-range NOEs.

basis of the above distance constraints, 37 dihedral angle constraints, and 30 hydrogen bond distance constraints. A final set of 24 structures was selected on the basis of minimum total energy, which had no violations of distance restraints ≥ 0.5 Å and dihedral angle restraints ≥ 5.0°. The ensemble coordinates of Gln25-RNase $T_1$ were deposited in the Brookhaven Protein Data Bank under the file name 1IYY. Structural statistics for the 24 converged structures are given in Table 1. Small root-mean-square (RMS) deviations of the bonds from idealized geometry (typically < 0.007 Å), as well as small values of the X-PLOR energies, indicated good geometry of the final set of the structures.

The average RMS differences (± standard deviation) of this ensemble were 0.85 ± 0.14 Å for the backbone atoms and 1.21 ± 0.11 Å for all heavy atoms (Table 1). Figure 1 shows the number of distance constraints at each residue, and the distribution of the RMS differences from the mean structure for the backbone atoms and all heavy atoms is shown in Figure 2A. For residues 2–5, 10–33, 38–42, 52–63, 76–92, and 100–103, the values of the RMS differences are smaller than 0.8 Å for the backbone atoms, and for residues 2–4, 10–25, 27, 30, 32–33, 38–41, 52–63, 74, 76, 78–80, 86–92, 101, and 103, they are smaller than 2.0 Å for all heavy atoms. These residues composed one α-helix and seven β-strands (Figure 2A).

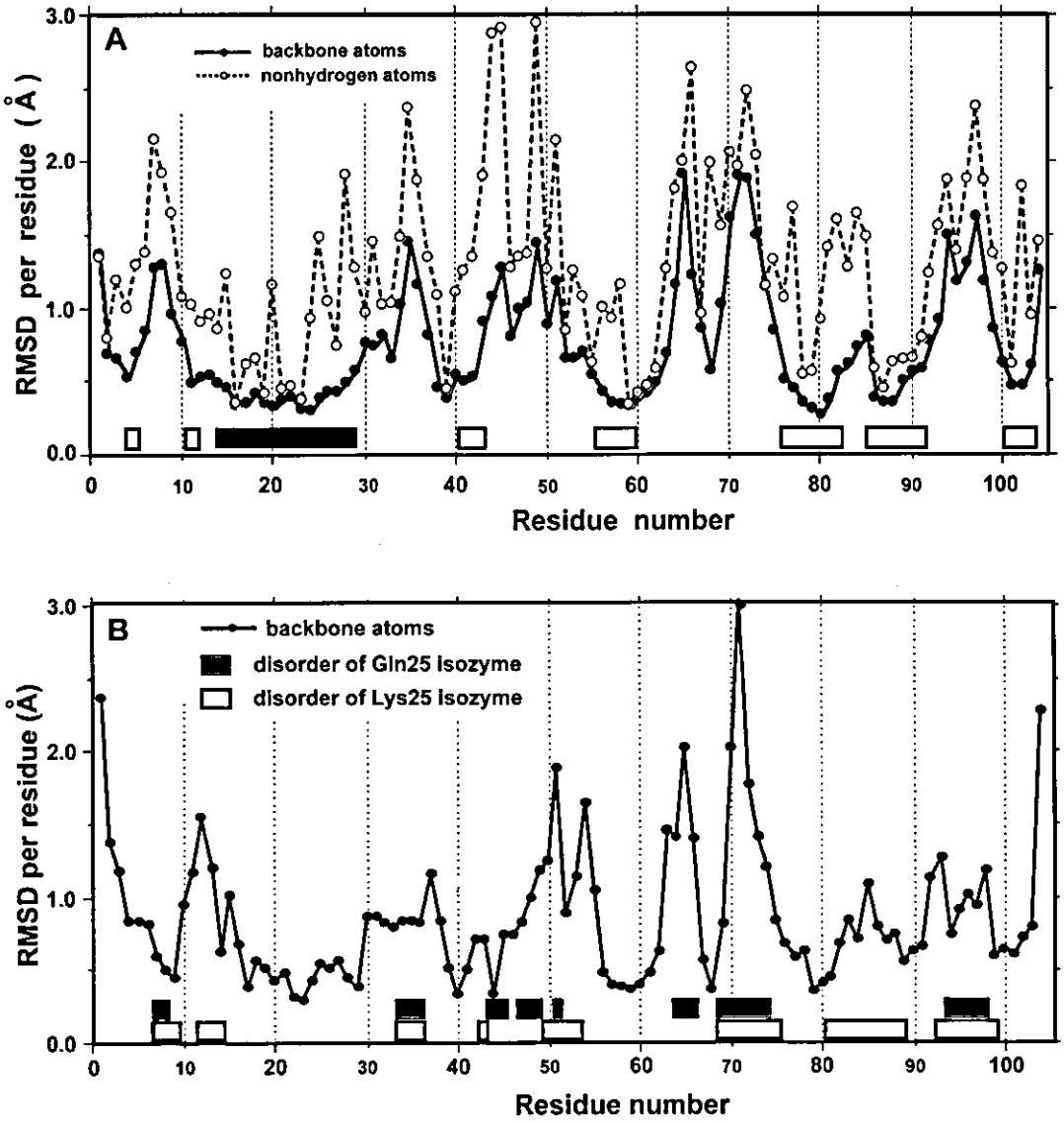


**Fig. 2** Distribution of the RMS Differences from the Mean Structure of Gln25-RNase $T_1$ (A) and between the Averaged Solution Structures of the Gln25- and Lys25-Isozymes (B).
(A) The averaged values of backbone atoms (N, $C_\alpha$, and C') and all non-hydrogen atoms were plotted as filled and open circles, respectively. A shaded or filled bar indicates the position of an α-helix or a β-strand, respectively. (B) The averaged value of backbone atoms (N, $C_\alpha$, and C') was plotted as filled circles. Filled and shaded bars indicate the diffused disorder segments of the Gln25 and Lys25 isozymes, respectively.

The individual backbone conformations of all residues in the final XPLOR ensemble are presented in the Ramachandran maps (Figure 3A and 3B). According to the analysis with the program PROCHECK, 52.5% of the residues are located in the most-favored regions, 45.4% in the allowed regions, and 2.1% in the disallowed regions. The majority of residues found in energetically unfavorable regions for non-glycine/proline residues were glycine residues (Figure 3A). However, most of the glycines were found in allowed regions for glycine (Figure 3B). Nevertheless, the plots of Gly7 (13 out of 24 structures) and Gly47 (20 out of 24 structures) existed in the disallowed region. These glycines were found to be located in the disordered loop regions (Figure 2A).

### Disulfide Bridge Geometry

As described in the Introduction, Pfeiffer and coauthors (1997) determined the solution structure of Lys25-RNase $T_1$ (PDB code 1YGW). Interestingly, they revealed that there are two conformations in solution for the torsion ($\chi_3$) angle of the disulfide bond between Cys2 and Cys10. In this study, we also analyzed all the $\chi_3$ angles in the 1IYY ensemble and tabulated the results as well as those of other solution and crystal structures of RNase $T_1$ (Table 2). In the 1IYY ensemble, there were two conformations for the $\chi_3$ angles, not only of the Cys2-Cys10 but also of the Cys6-Cys103. However, the major part of the chirality obtained for the Cys2-Cys10 and the Cys6-Cys103 adopted

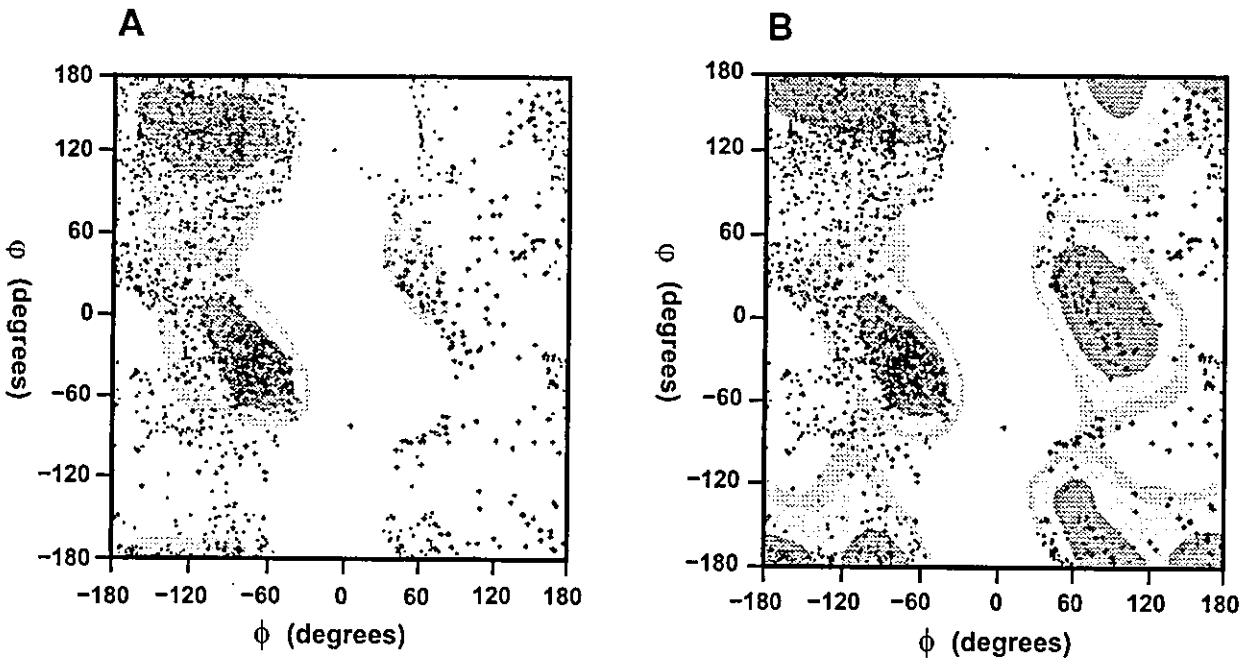


**Fig. 3** Ramachandran Plot for the 24 Final Structures of Gln25-RNase $T_1$.
Glycine residues are plotted with '+'. The grey-shaded backgrounds indicate allowed regions for non-glycine/proline residues (A) or glycine (B), which were determined from 378 different crystal structures with a resolution of 2.5 Å or better. The regions where the density of points ($\phi$, $\psi$) from this statistic is higher are shown by a shadow.

**Table 2** Comparison of the Cysteine Side-Chain Conformation in the Solution and Crystal Structures between the Two RNase $T_1$ Isozymes.

| | $\chi_3$ (degrees)[a] | | | |
|---|---|---|---|---|
| Isozyme | Gln25-RNase $T_1$ | | Lys25-RNase $T_1$ | |
| State | Solution[b] | Crystal[c] | Solution[b] | Crystal[c] |
| Cys2-Cys10 | −90±14 (18)/66±6 (6) | −114 (1RLS), −123 (1B2M) | −72±1 (16)/72±7 (18) | −115 (9RNT), −120 (1RNT) |
| Cys6-Cys103 | 119±30 (17)/−125±22 (7) | 114 (1RLS), 116 (1B2M) | 49±14 (34) | 110 (9RNT), 110 (1RNT) |

[a]Values corresponding to conformations near right-handed or left-handed disulfide bridges with $\chi_3 = 90°$ or $\chi_3 = -90°$, respectively, were averaged individually.
[b]Average $\chi_3$ angles of final XPLOR and DIANA ensembles comprising 24 and 34 conformers, respectively. The number of conformers of each group is given in parentheses.
[c]The PDB codes for the crystal structures are given in parentheses.

left-handed and right-handed conformations, respectively, which were consistent with those of the crystal structures (PDB codes 1RLS, 1B2M, 9RNT, and 1RNT). In solution, the chirality of the Cys6-Cys103 in Lys25-RNase $T_1$ adopted a right-handed conformation; on the other hand, that in the Gln25-isozyme showed a mixed chirality. For the Cys2, Cys10, and Cys103 residues in the Gln25-isozyme, the tight $\chi_1$ torsion-angle restraints were obtained from the quantitative analysis of the $\chi_1$ related $^3J$ couplings. Moreover, the β-methylene protons of the three cysteines were assigned stereospecifically. Therefore, the mixed chirality of the two disulfide bridges in the 1IYY ensemble is thought to be a genuine property in solution and not to be caused by a lack of the NOE data.

### Description of the Solution Structure of Gln25-RNase $T_1$

Figure 4 shows the backbone of the ensemble representing the solution structure of Gln25-RNase $T_1$. The region of the residues 13–29 (H-I) formed an α-helix with the typical hydrogen-bonding patterns including residues 12 to 30. The connectivity of $d_{\alpha N}(i,i+4)$ was observed from Ser12 to Glu28 intermittently; whereas $d_{\alpha N}(i,i+2)$ connectivity was observed only between Ser12 and Ser14, between Gly23 and Gln25, and between Glu28 and Gly30 (Kojima *et al.*, 1995). Accordingly, this indicates that this helix is an ordinary α-helix rather than a $3_{10}$-helix, consistent with the crystal structure of the Gln25-isozyme (Heinemann and Saenger, 1982; Sugio *et al.*, 1988; Arni *et al.*, 1992). On the other hand, in the Lys25-isozyme, the C-terminal residues of the H-I were more characteristic of a $3_{10}$-helix (Pfeiffer *et al.*, 1997). Furthermore, the side chains of Asp3, Ser12, and Ser13 in the Lys25-isozyme formed partially populated hydrogen bonds with the backbone amide protons of residues 13 and 14; however, such a feature (so-called 'N-capping' helix) was not observed in the Gln25-isozyme.

The seven segments of the residues 4–5 (B-I), 10–11 (B-II), 40–42 (B-III), 56–59 (B-IV), 76–81 (B-V), 86–92 (B-VI), and 100–103 (B-VII) were identified as β-strands. As shown in Figure 4, these β-strands are deduced to form two anti-parallel β-sheets as based on the inter-strand NOE connectivities and the results of the hydrogen-deuterium exchange experiments (Kojima *et al.*, 1995). One is located near the N-terminus of this protein, and is composed of the two short strands B-II and B-I; the remaining strands form the central anti-parallel five-stranded β-sheet (Figure 4). The hydrogen-bonding pattern between the backbone segments of the residues 76–81 and 86–92 is disturbed by insertion of residues in the latter strand, the residue 79 forming a classical β-bulge with the residues 87–88 (Kojima *et al.*, 1995). This is also observed in the solution structure of the Lys25-isozyme (Pfeiffer *et al.*, 1997). While the locations of the strands B-II, B-V, B-VI, B-VII and B-I agreed well with those of Lys25-RNase $T_1$, the strands B-III and B-IV had different locations: in the Lys25-isozyme, the strands B-III and B-IV comprised residues 38–40 and 57–61, respectively. In contrast to the solution structure, the N-terminal β-strands B-II and B-I in crystals were isolated from each other and did not form a β-sheet (Heinemann and Saenger, 1982; Sugio *et al.*, 1988; Arni *et al.*, 1992).

On the other hand, conformational disorder was observed in many loop regions (Figure 2A): the residues 6–9 (L-I), 30–39 (L-II), 43–55 (L-III), 60–75 (L-IV), and 93–99 (L-VI), except for residues 82–85 (L-V). The L-II, L-IV, and L-VI were located on the surface of the molecule facing toward the substrate and were in close vicinity (Figure 4). Their imprecise definition was most likely due to a lack of data caused by their high glycine content (Figure 1), since glycine residues cause a reduction of

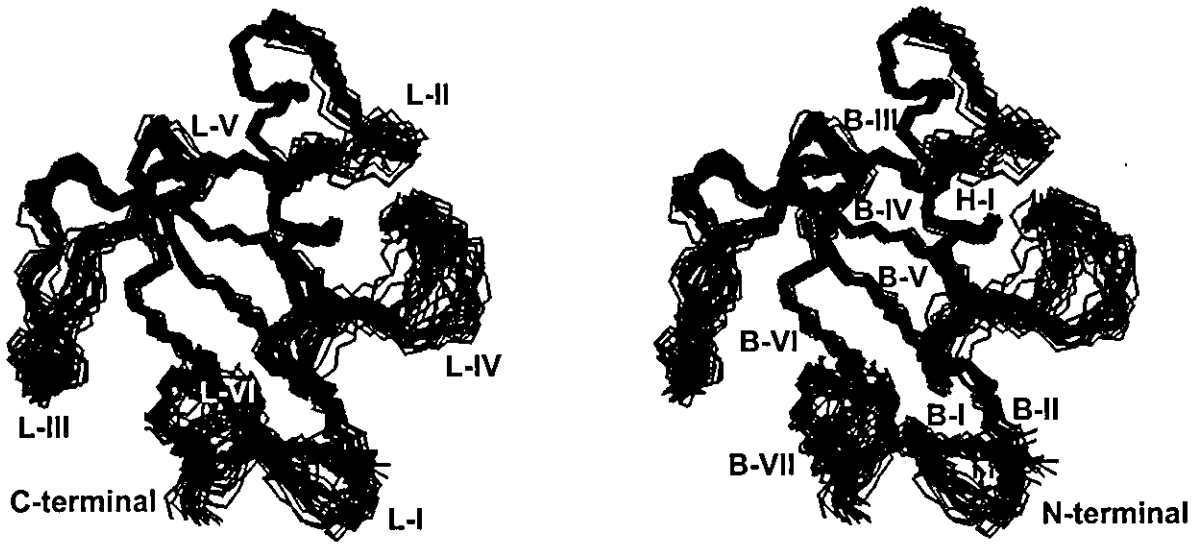


**Fig. 4** Stereo View of the Backbone of the Solution Structure of Gln25-RNase $T_1$.
The superposition was carried out so as to minimize the backbone RMS deviation from the mean structure for residues 2–103. The α-helix, seven β-strands, and five loops are labeled H-I, B-I to B-VII (the right view), and L-I to L-VI (the left view), respectively.

NOE data due to a low atom-packing density. On the other hand, the L-III, including one of the guanine-recognition sites, showed a high conformational disorder including all the heavy atoms (Figure 2A). The flexibility of the base-binding site has been suggested by X-ray crystallographic and NMR studies of the Lys25-isozyme (Martinez-Oyanedel *et al.*, 1991; Pfeiffer *et al.*, 1997). Therefore, the mobility of the L-III, including the base-binding site, may be essential to bind the guanine base of a substrate during catalysis.

### Comparison between the Calculated and Observed p$K_a$ Values of the Catalytically Important Residues

The p$K_{calc}$ values of the catalytically important residues (His40, Glu58, and His92) and His27 of the 1IYY ensemble are listed in Table 3 as well as the corresponding p$K_{exp}$ values determined previously (Inagaki *et al.*, 1981; Spitzner *et al.*, 2001). The p$K_{calc}$ values of the His27 imidazole group and the Glu58 carboxylate agreed comparatively well with the corresponding p$K_{exp}$ values. On the other hand, the p$K_{calc}$ values of the His40 and His92 imidazole groups were in poor agreement with the corresponding p$K_{exp}$ values. Moreover, the p$K_{calc}$ values of His40, Glu58, and His92 had large standard deviations (more than 2 pH units). Therefore, we have classified the NMR structures into groups on the basis of the p$K_{calc}$ values of the His40 imidazole, Glu58 carboxylate, and His92 imidazole groups in order to consider the real microenvironment of each ionizable group (Hatano *et al.*, 2003).

With respect to the His40 imidazole group, the 24 structures were classified into two groups in terms of p$K_{calc}$. One (models #1, 2, 4–6, 15, 18, and 24 in the 1IYY ensemble) is a group with a low p$K_{calc}$ value (average 3.8±2.0), whereas the other group (all the remainder in 1IYY) has a high p$K_{calc}$ value (7.3±0.6). The former group has a low value (−2.1±1.5) of $\Delta pK_{interact}$, and the His40 imidazole group is close to the side chains of Tyr38 and

**Table 3** Ionization Constants of the Catalytic Groups in the Ensemble of Gln25-RNase $T_1$.

| Ionizable group | p$K_{model}$[a] | p$K_{calc}$ | Solvent accessibility[b] | p$K_{exp}$[c] |
|---|---|---|---|---|
| His27[d] | 7.0 | 7.6±0.7 | 56±8 | 7.3±0.0 (7.0±0.0) |
| His40 | 7.0 | 6.1±2.1 | 71±6 | 7.9±0.0 (7.8±0.1) |
| Glu58 | 4.4 | 3.5±2.2 | 22±8 | 4.3±0.1 (4.0±0.0) |
| His92 | 7.0 | 2.6±2.0 | 38±10 | 7.8±0.0 (7.3±0.0) |

[a]Model compound p$K_a$ values used are taken from data on the titration behavior of denatured proteins (Nozaki and Tanford, 1967).
[b]Solvent-accessible surface areas per residue (Å$^2$) were estimated from the 1IYY ensemble using the program MOLMOL (Koradi *et al.*, 1996) and a probe size of 1.4 Å.
[c]These values of Gln25-RNase $T_1$ were quoted from (Inagaki *et al.*, 1981). The corresponding values of Lys25-RNase $T_1$ are also given in parentheses (Spitzner *et al.*, 2001).
[d]This residue is not one of the catalytic residues, but is indicated in the Table for comparison.

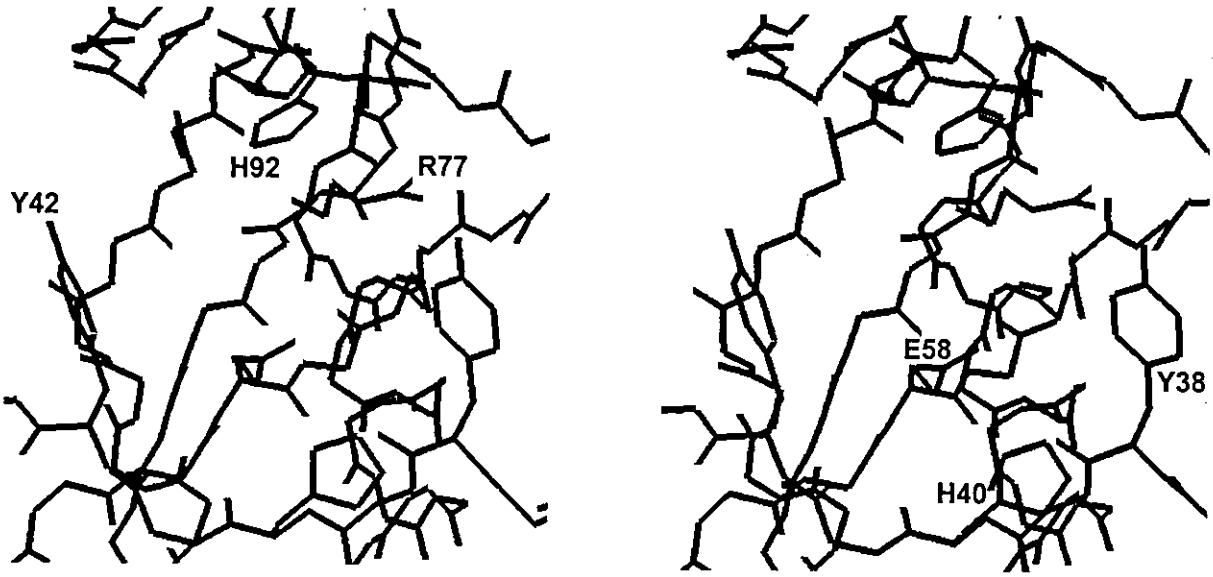


**Fig. 5** Stereo View of the Catalytic Site of Gln25-RNase $T_1$ in Solution.
The backbone of the average structure and the side chains of the catalytic residues are shown. The catalytic and other important residues are labeled.

Arg77 in space; on the other hand, the latter has a large $\Delta pK_{interact}$ value (1.1 ± 0.8). It is likely that the proton of the imidazole group in the former group can be dissociated easily due to the hydrophobic interaction with the aromatic ring of Tyr38 and electrostatic repulsion by the Arg77 guanidino group. Considering the $pK_{exp}$ value of His40, this residue in reality would be away from the side chains of Tyr38 and Arg77 as in the latter group (Figure 5).

In the case of the Glu58 carboxylate, we classified the 24 structures into two groups, one group (models #1, 2, 4, 15, and 21) with a high $pK_{calc}$ value (the average, 7.0 ± 0.8) and the other with a low $pK_{calc}$ value (2.5 ± 1.3). The former group has a smaller solvent-accessible area (average 11 ± 3 Å$^2$), with the carboxyl group prevented from the solvent access by the aromatic rings of Tyr42 and Phe100, as compared with that of the other group (25 ± 7 Å$^2$). Taking the $pK_{exp}$ value of Glu58 into consideration, this residue would not be prevented from such solvent access in reality. In addition, the site-site interaction potential analysis based on the $\Delta pK_{interact}$ values (Figure 6) indicated that the ionization of the carboxyl group of Glu58 is influenced strongly by the nearby His40 imidazole and Arg77 guanidino groups. Therefore, such an electrostatic interaction with the positive charges of His40 and Arg77 may contribute in a major way to depression of the $pK_a$ value of the carboxyl group of Glu58.

As for the imidazole group of His92, the ensemble was classified into two groups; one group (model #2, 4, 5, 8, 11, 15, 17, 18, and 20–23) with a low $pK_{calc}$ value (average 1.2 ± 1.6) and the other with a high $pK_{calc}$ value (average 4.3 ± 0.8). The former group, of which the imidazole group has strong interactions with the titratable groups of Tyr42 and Arg77 (Figure 6), has a larger $\Delta pK_{interact}$ value (−2.9 ± 0.6) than that of the other group (−1.0 ± 0.6). The averaged $pK_{calc}$ value of the former group indicates that the His92 imidazole group can dissociate readily due to a hydrophobic interaction with the aromatic ring of Tyr42 and electrostatic repulsion by the Arg77 guanidino group. Considering the $pK_{exp}$ value of His92, in reality it would be pointing away from the side chains of Tyr42 and Arg77 as in the latter group (Figure 5). Nevertheless, the $pK_{calc}$ value of the ionizable group of His92 is markedly low as compared with that of the model compound (Table 3). The difference (4.4) between the $pK_{calc}$ and $pK_{model}$ values consisted of $\Delta pK_{Born}$ of −2.8 ± 0.5, $\Delta pK_{back}$ of 0.4 ± 0.9, and $\Delta pK_{interact}$ of −2.0 ± 1.1. Among these contributions, the Born term was notable as well as the site-site interaction term. Probably most of the side chain of His92 would be buried inside the molecule as judged from its low solvent accessibility (Table 3). The former group has a larger $\Delta pK_{Born}$ value (the average, −3.2 ± 0.2), where the imidazole group has a hydrophobic interaction with the Tyr38 and Tyr42 rings (Figures 5 and 6), as compared with that of the latter (−2.4 ± 0.3). Therefore, the discrepancy in $pK_{calc}$ value for the His92 imidazole group is thought to be due to the large $\Delta pK_{Born}$ value as well as the $\Delta pK_{interact}$ value.

As for the His92 imidazole group, Inagaki and coauthors (1981) described that the remarkable pH dependence of the C2-proton chemical shift was observed in Gln25-RNase $T_1$. The chemical shift of the C2 proton was abnormally up-field (~7.5 ppm) at pH 6 and then showed two significant inflections with extrinsic $pK_a$'s of 4.2 ± 0.1 and 1.0 ± 0.2, which indicates some interaction with a carboxylate-bearing residue. However, from the site-site interaction analysis of Gln25-RNase $T_1$, such a carboxylate-bearing residue was not found anywhere (Figure 6). Therefore, we could not exclude the possibility that the two significant inflections are involved in the $pK_{calc}$ value of His92. In summary, models #3, 7, 9, 13, and 14 in the 1IYY ensemble were revealed to satisfy all the obtained $pK_{exp}$ values.

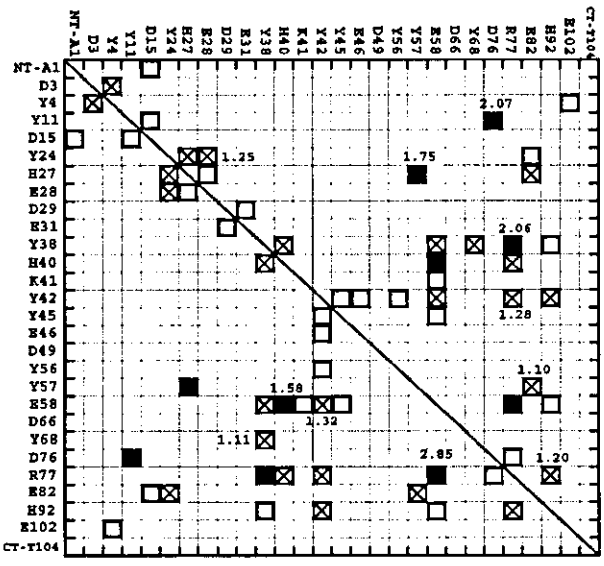


**Fig. 6** Diagonal Plot Representation of the Site-Site Interaction Analysis Based on the Matrices of $\Delta pK_{interact}$ Calculated for the 1IYY Ensemble.

The two axes are scaled with the ionizable groups of Gln25-RNase $T_1$. The squares connect the pairs of the ionizable groups for which site-site interactions were observed: filled, strong ($\Delta pK_{interact} \geq 1.4$); hatched, medium ($0.7 \leq \Delta pK_{interact} < 1.4$); open squares, weak interactions ($0.1 \leq \Delta pK_{interact} < 0.7$). The averaged $\Delta pK_{interact}$ values higher than 1.0 are displayed beside the squares.

## Discussion

Both isozymes of RNase $T_1$ are common in nature. The difference in the coding sequence of the two isozymes at amino acid 25 is an A→C nucleotide transversion, resulting in the change in the codon from lysine to glutamine. The only other isozyme that exists with a nucleotide point mutation at this position (A→G) is Glu25-RNase $T_1$. However, this mutant is apparently deleterious, since it has not been detected in nature. At the outset of this study, we had interesting questions about whether there is any structural or functional evolutionary selection pressure that would favor one isozyme over the other under any circumstances.

In addition to the NMR studies and X-ray crystallography studies of this intensively investigated model enzyme, there are also other significant studies. A detailed functional study (Heinemann *et al.*, 1980) reported no differences in catalytic activity or substrate specificity between the two isozymes. This is not surprising since amino acid 25, located in the α-helix of RNase $T_1$, is on the opposite side of the molecule from the active site. Other important studies include a number of stability studies on the two isozymes performed by chemical or heat denaturation (Shirley *et al.*, 1989; Kiefhaber *et al.*, 1990, Yu *et al.*, 1994). The Gln25-isozyme was found to be less stable by about 0.9 kcal. In their study on the stability of Lys25-RNase $T_1$, Kiefhaber and coauthors (1990) rationalized the differences in stability by energy minimization calculations, suggesting that a salt bridge occurs between Glu28 and Lys25 in Lys25-RNase $T_1$ that cannot occur in the Gln25-isozyme. Their studies were done in 1990 when only the X-ray structure had been published (Heinemann and Saenger, 1982).

Here we addressed this stability issue by comparing the NMR solution structures of both isozymes. Indeed, no significant differences were reported in the X-ray crystallography studies so that a thorough comparison of the NMR solution structures and comparison of these with the X-ray structures would be welcomed in addressing the stability issue. As described in the Results section, we revealed that Gln25-RNase $T_1$ has neither an N-capping helix nor a partial $3_{10}$-helix in the helical end as compared with the Lys25-isozyme. Furthermore, we have calculated the averaged solution structures of the two isozymes and a comparison between them revealed that backbone RMS differences of the residues 11–13 (B-II), 37 (L-II), 49–51 (L-III), 53–55 (L-III), 63–66 (L-IV), 69–74 (L-IV), 85 (L-V), 92–93 (L-VI), and 98 (L-VI) were more than 1.0 Å (Figure 2B). In Lys25-RNase $T_1$, the conformational disorder was shown to present in the backbone segments of residues 7–9, 12–14, 34–36, 43, 50–53, 69–75, 81–88, and 93–99, in which backbone RMS deviations were more than 1.0 Å (Pfeiffer *et al.*, 1997; Figure 2B). On the other hand, the backbone disorder in the Gln25-isozyme existed in the residues 7–8, 34–36, 44–45, 47–49, 51, 64–66, 69–74, and 94–98, in which backbone RMS differences were more than 1.0 Å (Figure 2A, B). Therefore, the conformations of the residues 11–13, 63–66, and 92–93 are thought to be sensitive for the replacement of Lys25 with Gln.

In the previous study, the replacement of Lys25 with Gln was shown to definitely affect the NH and/or $C_\alpha H$ chemical shifts of Lys/Gln25 (the values shifted, 0.15 ppm/0.12 ppm), Leu26 (0.15 ppm/0 ppm), Asp29 (0.23 ppm/0.10 ppm), Phe48 (0.16 ppm/0.05 ppm), and Thr93 (0.02 ppm/0.10 ppm) (Kojima *et al.*, 1995). Furthermore, the NH and/or $C_\alpha H$ chemical shifts in the Gln25 isozyme were also investigated under another experimental condition at 305 K and pH 4.4 (Kojima *et al.*, unpublished results). As compared with the chemical shifts of the Gln25-isozyme under the conditions of this study, the NH and/or $C_\alpha H$ chemical shifts of the residues Thr5 (values shifted by 0.25 ppm/0.01 ppm), Ser12 (0.28 ppm/0.08 ppm), Ser13 (0.19 ppm/0.01 ppm), Asp29 (0.07 ppm/0.01 ppm), Asn44 (0.21 ppm/0.02 ppm), Glu46 (0.12 ppm/0.01 ppm), and Cys103 (0.08 ppm/0.06 ppm) were affected by 0.05 ppm or more. These results indicate that the residues Asp29 and Thr93 are sensitive for the Lys25-to-Gln replacement and that the residues Thr5, Ser12, Ser13, and Cys103 are sensitive for the thermal and/or pH change. As described in the Results section, the base-binding site, including Tyr42 through Glu46 and Asn98, has been suggested to be very flexible by an X-ray crystallographic study; therefore, the chemical shifts of Asn44 and Glu46 would be easily affected by the experimental conditions. On the other hand, it is noteworthy that the residues 11–13 (B-II) and 92–93 (L-VI) exist in the well-converged regions (Figure 2A).

From a molecular dynamics calculation, Suzuki and coauthors (1995) concluded that Asp29 and/or Glu31, but not Glu28, might be the counterpart of the electrostatic interaction of Lys25 in Lys25-RNase $T_1$. This is consistent with the results that the $C_\beta H_2$ chemical shift values of Asp29 and Glu31 for the Gln25-isozyme differ by 0.27 ppm from those for the Lys25-isozyme as well as the NH and $C_\alpha H$ chemical shifts (Kojima *et al.*, 1995). Indeed, the side chain of Lys/Gln25 was closest to that of Asp29 in the NMR structures of 1YGW and 1IYY ensembles (Table 4), whose conformation would be more favorable both structurally and energetically. However, in Gln25-RNase $T_1$, a salt bridge between Asp29/Glu31 and Gln25 could not be formed due to the lack of the electrostatic interaction. Probably the lack of such a salt bridge would induce the conformational perturbation of the N-terminal end of the α-helix in the Gln25-isozyme, leading to the loss of an 'N-capping helix' structure as reported in the Lys25-isozyme. This is consistent with the result that the RMS differences of the residues 11–13 between both isozymes were more than 1.0 Å (Figure 2B). Furthermore,

**Table 4** Distances between the Specific Atoms in Each Ensemble of RNase $T_1$.

| Ensemble name<br>Starting point | | Gln25-RNase $T_1$ (1IYY)<br>NE2 atom of Gln25 | Lys25-RNase $T_1$ (1YGW)<br>NZ atom of Lys25 |
|---|---|---|---|
| Destination | CD atom of Glu28 | 9.1±1.7 Å | 6.9±1.1 Å |
| | CG atom of Asp29 | 7.0±1.6 Å | 5.5±1.3 |

the perturbation of this segment might transfer to the residues 92–93 *via* the two disulfide bonds, which reflects the differences in torsion angles of the disulfide bonds between these isozymes. The segment of residues 11–13 is also very sensitive for thermal and/or pH change in Gln25-RNase $T_1$. It is most likely that this perturbation transfers to the disulfide bond Cys6-Cys103, which is consistent with the above-mentioned differences of the backbone chemical shifts.

In conclusion, Gln25-RNase $T_1$ is revealed to have no salt bridge between Asp29/Glu31 and Gln25, so that the helical region losing a salt bridge would perturb the N-terminal helical end and this perturbation might transfer to the residues 92–93 *via* the two disulfide bonds. These perturbations would be involved in the stability of Lys25-RNase $T_1$ to the Gln25-isozyme. On the other hand, the negative charge of Glu25 in the Glu25-RNase $T_1$ would receive an electrostatic repulsion from the negative-charge cluster including Glu28, Asp29, and Glu31; hence, the above-mentioned perturbations in Glu25-RNase $T_1$ would be more serious than those in the Gln25-isozyme. Accordingly, it is most likely that structurally evolutionary pressure favors the Lys25- and Gln25-isozyme more than the other ones, especially the Glu25-RNase $T_1$. It is very interesting to examine the 3-D structure, catalytic activity, and thermodynamic stability of Glu25-RNase $T_1$.

## Materials and Methods

### 2-D NMR Measurement

The sample was dissolved at 2 mM in 10% $D_2O$/90% $H_2O$, pH 5.5. Nuclear Overhauser effect spectroscopy (NOESY) (Jeener *et al.*, 1979; Kumar *et al.*, 1980) spectra were recorded on a Bruker AMX600 spectrometer at 313 K, and the chemical-shift values were referenced to the external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). The spectra were recorded in a time-proportional phase-incrementation mode (Drobny *et al.*, 1979; Bodenhausen *et al.*, 1980). The chemical shifts of individual protons of the NOESY spectra were assigned referring to the previous results (Kojima *et al.*, 1995).

### Experimental Constraints for Structure Calculation

We selected the NOESY spectra with a mixing time of 150 ms for structure analysis, since it provided the largest number of NOE cross-peaks and the best signal-to-noise ratio. The cross peaks were picked up using the software package Felix (Biosym Technologies, Inc., San Diego, USA). Taking into account the two-spin diffusion and internal motion, all the observable NOE intensities were estimated using the program CORMA (Keepers and James, 1984). The NOE information was classified as 1.8–2.7, 1.8–3.5, or 1.8–5.0 Å on the basis of the intensity. Appropriate pseudo-atom corrections were applied to non-stereospecifically assigned protons (Wüthrich, 1984).

The fifteen hydrogen-bonded amides (Kojima *et al.*, 1995) were defined as two additional distance constraints, 2.5 Å $\leq N_{(i)}$-$O_{(i)} \leq$ 3.5 Å and 1.5 Å $\leq NH_{(i)}$-$O_{(i)} \leq$ 2.5 Å. Constraints on the $\phi$ angle were classified according to the size of the $^3J_{HNH\alpha}$ coupling constant (Pardi *et al.*, 1984). For the 21 residues, of which $^3J_{HNH\alpha}$ were more than 8 Hz, the $\phi$ angles were constrained to the range $-160° \leq \phi \leq -80°$. Constraints on $\chi_1$ were determined by examination of DQF-COSY and NOESY spectra (Wagner *et al.*, 1987). From this analysis, one was constrained to $20° \leq \chi_1 \leq 100°$, eight were constrained to $-100° \leq \chi_1 \leq -20°$, and five were constrained to $140° \leq \chi_1 \leq 220°$. The *cis* $\omega$ angles for the two X-Pro peptide linkages were constrained to 0° ± 80° (Stewart *et al.*, 1990). The three explicit disulfide bonds were constrained to $C_{\beta(i)}$-$S_{\gamma(i)}$ = 2.99 Å ± 0.50 Å and $S_{\gamma(i)}$-$S_{\gamma(i)}$ = 2.02 Å ± 0.02 Å.

### Structure Calculation

On the basis of the interproton distances and dihedral angle restraints, structure calculations were performed with the program X-PLOR version 3.1 (Brünger, 1992). Calculations were started from random coil conformations as initial structures and the structures were optimized using the hybrid distance geometry/simulated annealing protocol. Several rounds of structure calculation and assignment were performed in order to resolve the ambiguities. The final values of the force constants used are as follows: 50 kcal $mol^{-1}$ $Å^{-2}$ and 200 kcal $mol^{-1}$ $rad^{-2}$ for the square-well interproton distance and torsion angle constraints, respectively; and 4 kcal $mol^{-1}$ $Å^{-4}$ for the quartic van der Waals repulsion term with the hard-sphere van der Waals radii set to 0.8 times their values in the energy parameters of the 'parallhdg.pro' file (Brooks *et al.*, 1983).

The resulting conformation ensemble was analyzed with the program PROCHECK (Laskowski *et al.*, 1996). The solvent-accessible surface area per residue was analyzed using the program MOLMOL version 2k.1 (Koradi *et al.*, 1996). Computer graphic representation was obtained using the software package of MidasPlus (Ferrin *et al.*, 1988). All calculations and graphical displays were performed on IRIS $O_2$ color graphics workstations (Silicon Graphics Inc.).

### $pK_a$ Calculation

The $pK_a$ values of the ionizable groups were calculated with the program MEAD (Bashford and Gerwert, 1992; Bashford, 1997), based on the atomic coordinates of the molecular structures determined (PDB codes, 1IYY and 1YGW). The program allowed estimation of the three types of electrostatic contributions to the titrating group by solving a linearized Poisson-Boltzmann equation both in the protein molecule and in model compounds (*N*-formyl *N*-methyl amide derivatives). The three contributions are: $\Delta\Delta G_{Born}$, the interaction of the titrating charge with the polarization that this charge itself induces in the surroundings; $\Delta\Delta G_{back}$, the interaction of the titrating charge with the non-titrating (background) charges such as a peptide dipole; and $\Delta\Delta G_{interact}$, the interaction of the titrating charge with other ionizable charges (Bashford and Karplus, 1990). Calculations were performed with an internal dielectric constant of 4, a dielectric constant of the surrounding water of 80, an ionic strength of 0.1, temperature of 313 K, and others including lattice settings of default parameter values. The background partial charge of each atom was taken from the AMBER all-atom force-field parameters (Weiner *et al.*, 1986), and the atomic radii were the standard van der Waals radii (Bondi, 1964). $\Delta pK_{Born}$, $\Delta pK_{back}$, and $\Delta pK_{interact}$ are defined as the corresponding $pK_a$ shift due to the $\Delta\Delta G_{Born}$, $\Delta\Delta G_{back}$, and $\Delta\Delta G_{interact}$, respectively as:

$$\Delta pK_{Born} = \Delta\Delta G_{Born}/(2.303\, k_B T)$$
$$\Delta pK_{back} = \Delta\Delta G_{back}/(2.303\, k_B T)$$
$$\Delta pK_{interact} = \Delta\Delta G_{interact}/(2.303\, k_B T)$$

where $k_B$ and $T$ represent the Boltzmann constant and the absolute temperature, respectively. $pK_{calc}$ was estimated due to the total contribution of these three $pK_a$ shifts as:

$pK_{calc} = pK_{model} - (\Delta pK_{Born} + \Delta pK_{back} + \Delta pK_{interact})$

where $pK_{model}$ is the $pK_a$ of a corresponding model compound (Nozaki and Tanford, 1967).

## Acknowledgments

We are very grateful to Dr. Donald Bashford (Scripps Research Institute, USA) for the use of the MEAD program. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan. We also thank Dr. Mitsugu Suzuki (The University of Electro-Communications) for the help of semi-automated NOE assignments.

## References

Arata, Y., Kimura, S., Matsuo, H., and Narita, K. (1979). Proton and phosphorus nuclear magnetic resonance studies of ribonuclease $T_1$. Biochemistry *18*, 18–24.

Arni, R., Heinemann, U., Tokuoka, R., and Saenger, W. (1988). Three-dimensional structure of the ribonuclease $T_1$-2'GMP complex at 1.9 Å resolution. J. Biol. Chem. *263*, 15358–15368.

Arni, R.K., Pal, G.P., Ravichandran, K.G., Tulinsky, A., Walz Jr., F.G., and Metcalf, P. (1992). Three-dimensional structure of Gln25-ribonuclease $T_1$ at 1.84-Å resolution: structural variations at the base recognition and catalytic sites. Biochemistry *31*, 3126–3135.

Bashford, D. (1997). In: Scientific computing in object-oriented parallel environments. Comput. Sci. *1343*, 233–240.

Bashford, D. and Karplus, M. (1990). $pK_a$'s of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry *29*, 10219–10225.

Bashford, D. and Gerwert, K. (1992). Electrostatic calculations of the $pK_a$ values of ionizable groups in bacteriorhodopsin. J. Mol. Biol. *224*, 473–486.

Bodenhausen, G., Vold, R.L., and Vold, R.R. (1980). Multiple quantum spin-echo spectroscopy. J. Magn. Reson. *37*, 93–106.

Bondi, A. (1964). van der Waals volumes and radii. J. Chem. Phys. *65*, 441.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Saminathan, S., and Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculation. J. Comput. Chem. *4*, 187–217.

Brünger, A.T. (1992). X-PLOR Manual Version 3.1: A System for X-ray Crystallography and NMR (New Haven, USA: Yale University Press).

De Vos, S., Doumen, J., Langhorst, U., Steyaert, J. (1998). Dissecting histidine interactions of ribonuclease $T_1$ with asparagine and glutamine replacements: analysis of double mutant cycles at one position. J. Mol. Biol. *275*, 651–661.

Drobny, G.P., Pines, A., Sinton, S., Weitekamp, D.P., and Wemmer, D.E. (1979). Fourier transform multiple quantum nuclear magnetic resonance. Symp. Faraday Soc. *23*, 49–55.

Egami, F., Takahashi, K., and Uchida, T. (1964). Ribonucleases in Taka-Diastase: properties, chemical nature, and applications. Progr. Nucleic Acid Res. Mol. Biol. *3*, 59–101.

Ferrin, T.E., Huang, C.C., Jarvis, L.E., and Langridge, R. (1988). The MIDAS display system. J. Mol. Graphics *6*, 12–27.

Fushman, D., Weisemann, R., Thüring, H., and Rüterjans, H. (1994). Backbone dynamics of ribonuclease $T_1$ and its complex with 2'GMP studied by two-dimensional heteronuclear NMR spectroscopy. J. Biomol. NMR *4*, 61–78.

Gohda, K., Oka, K., Tomita, K., Hakoshima, T. (1994). Crystal structure of RNase $T_1$ complexed with the product nucleotide 3'-GMP. Structural evidence for direct interaction of histidine 40 and glutamic acid 58 with the 2'-hydroxyl group of the ribose. J. Biol. Chem. *269*, 17531–17536.

Hakoshima, T., Itoh, T., Tomita, K., Goda, K., Nishikawa, S., Morioka, H., Uesugi, S., Ohtsuka, E. and Ikehara, M. (1992). Three-dimensional structure of a mutant ribonuclease $T_1$ (Y45W) complexed with non-cognizable ribonucleotide, 2'AMP, and its comparison with a specific complex with 2'GMP. J. Mol. Biol. *223*, 1013–1028.

Hatano, K-i., Kojima, M., Tanokura, M., and Takahashi, K. (2003). Nuclear magnetic resonance studies on the $pK_a$ values and interactions of ionizable groups in bromelain inhibitor VI from pineapple stem. Biol. Chem. *384*, 83–91.

Heinemann, U., Wernitz, M., Pahler, A., Saenger, W., Menke, G., and Rüterjans, H. (1980). Crystallization of a complex between ribonuclease $T_1$ and 2'-guanylic acid. Eur. J. Biochem. *109*, 109–114.

Heinemann, U. and Saenger, W. (1982). Specific protein-nucleic acid recognition in ribonuclease $T_1$-2'-guanylic acid complex: an X-ray study. Nature *299*, 27–31.

Hoffmann, E. and Rüterjans, H. (1988). Two-dimensional $^1$H-NMR investigation of ribonuclease $T_1$. Resonance assignments, secondary and low-resolution tertiary structures of ribonuclease $T_1$. Eur. J. Biochem. *177*, 539–560.

Inagaki, F., Kawano, Y., Shimada, I., Takahashi, K., and Miyazawa, T. (1981). Nuclear magnetic resonance study on the microenvironments of histidine residues of ribonuclease $T_1$ and carboxymethylated ribonuclease $T_1$. J. Biochem. *89*, 1185–1195.

Jeener, J., Meiler, B.H., Bachmann, P., and Ernst, R.R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. J. Chem. Phys. *71*, 4546–4533.

Keepers, J.W. and James, L.T. (1984). A theoretical study of distance determinations from NMR. J. Magn. Reson. *57*, 404–426.

Kiefhaber, T., Schmid, F. X., Renner, M., Hinz, H.-J., Hahn, U., and Quaas, R. (1990). Stability of recombinant Lys25-ribonuclease $T_1$. Biochemistry *29*, 8250–8257.

Kojima, M., Miyano, H., Suzuki, E., Tanokura, M., Murakami, T., and Takahashi, K. (1995). Effects of replacement of Lys25 with Gln on the conformation of ribonuclease $T_1$: sequence-specific $^1$H NMR resonance assignments of Gln25 ribonuclease $T_1$ by two-dimensional NMR spectroscopy. J. Biochem. *118*, 710–716.

Kojima, M., Tanokura, M., and Takahashi, K. (2000). Structure-function relationship of ribonuclease $T_1$ based on molecular structure. Bioimages *8*, 45–55.

Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graphics *14*, 51–55.

Kumar, A., Ernst, R.R., and Wüthrich, K. (1980). A two-dimensional nuclear Overhauser enhancement experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem. Biophys. Res. Commun. *95*, 1–6.

Kyogoku, Y., Watanabe, M., Kainosho, M., and Oshima, T. (1982). A $^{15}$N-NMR study on ribonuclease $T_1$-guanylic acid complex. J. Biochem. *91*, 675–679.

Landt, O., Tholke, J., Grunert, H.P., Saenger, W., and Hahn, U. (1997). Ribonuclease $T_1$ is active when both catalytic histidines are replaced by aspartate. Biol. Chem. *378*, 553–558.

Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Loverix, S. and Steyaert, J. (2001). Deciphering the mechanism of RNase $T_1$. Methods Enzymol. *341*, 305–323.

Martinez-Oyanedel, J., Choe, H.-W., Heinemann, U., and Saenger, W. (1991). Ribonuclease $T_1$ with free recognition and catalytic site: crystal structure analysis at 1.5 Å resolution. J. Mol. Biol. *222*, 335–352.

McNutt, M., Mullins, L.S., Raushel, F.M., and Pace, C.N. (1990). Contribution of histidine residues to the conformational stability of ribonuclease $T_1$ and mutant Glu58/Ala. Biochemistry *29*, 7572–7576.

Nishikawa, S., Morioka, H., Kim, H. J., Fuchimura, K., Tanaka, T., Uesugi, S., Hakoshima, T., Tomita, K., Ohtsuka, E., and Ikehara, M. (1987). Two histidine residues are essential for ribonuclease $T_1$ activity as is the case for ribonuclease A. Biochemistry *26*, 8620–8624.

Nozaki, Y. and Tanford, C. (1967). Examination of titration behavior. Methods Enzymol. *11*, 715–734.

Pardi, A., Billeter, M., and Wüthrich, K. (1984). Calibration of the angular dependence of the amide proton-$C_\alpha$ proton coupling constants, $^3J_{HN\alpha}$, in a globular protein. Use of $^3J_{HN\alpha}$ for identification of helical secondary structure. J. Mol. Biol. *180*, 741–751.

Pfeiffer, S., Karimi-Nejad, Y. and Rüterjans, H. (1997). Limits of NMR structure determination using variable target function calculations: ribonuclease $T_1$, a case study. J. Mol. Biol. *266*, 400–423.

Shimada, I. and Inagaki, F. (1990). Binding modes of inhibitors of ribonuclease $T_1$ as elucidated by analysis of two-dimensional NMR. Biochemistry *29*, 757–764.

Shirley, B. A., Stanssens, P., Steyaert, J., and Pace, C. N. (1989). Conformational stability and activity of ribonuclease $T_1$ and mutants. J. Biol. Chem. *264*, 11621–11625.

Spitzner, N., Lohr, F., Pfeiffer, S., Koumanov, A., Karshikoff, A., Ruterjans, H. (2001). Ionization properties of titratable groups in ribonuclease $T_1$. I. p$K_a$ values in the native state determined by two-dimensional heteronuclear NMR spectroscopy. Eur. Biophys. J. *30*, 186–197.

Stewart, D.E., Sarker, A., and Wampler, J.E. (1990). Occurrence and role of *cis* peptide bonds in protein structures. J. Mol. Biol. *214*, 253–260.

Steyaert, J., Hallenga, K., Wyns, L., and Stansssens, P. (1990). Histidine-40 of ribonuclease $T_1$ acts as base catalyst when the true catalytic base, glutamic acid-58, is replaced by alanine. Biochemistry *29*, 9064–9072.

Steyaert, J., Opsomer, C., Wyns, L., and Stanssens, P. (1991). Quantitative Analysis of the contribution of glu46 and asn98 to the guanosine specificity of ribonuclease $T_1$. Biochemistry *30*, 494–499.

Steyaert, J. and Wyns, L. (1993). Functional interactions among the His40, Glu58 and His92 catalysts of ribonuclease $T_1$ as studied by double and triple mutants. J. Mol. Biol. *229*, 770–781.

Steyaert, J. (1997). A decade of protein engineering on ribonuclease $T_1$ - atomic dissection of the enzyme-substrate interactions. Eur. J. Biochem. *247*, 1–11.

Sugio, S., Amisaki, T., Ohishi, H., and Tomita, K. (1988). Refined X-ray structure of the low pH form of ribonuclease $T_1$-2'-guanylic acid complex at 1.9 Å resolution. J. Biochem. *103*, 354–366.

Suzuki, R., Kojima, M., and Tanokura, M. (1995). The relationship between thermodynamic stability and molecular structure of Lys25-ribonuclease $T_1$. Bioimages *3*, 65–69.

Takahashi, K. and Moore, S. (1982). Ribonuclease $T_1$, In: The Enzymes, Vol. 15, P. D. Boyer, ed. (New York, USA: Academic Press), pp.453–468.

Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Go, N., and Wüthrich, K. (1987). Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. J. Mol. Biol. *196*, 611–639.

Weiner, S.J., Kollman, P.A., Nguyen, D.T., and Case, D.A. (1986). An all atom force field for simulations of proteins and nucleic acids. J. Comput. Chem. *7*, 230–252.

Weisemann, R., Rüterjans, H., Schwalbe, H., Schleucher, J., Bermel, W., and Griesinger, C. (1994). Determination of $H^N$, $H_\alpha$ and $H^N$, C' coupling constants in $^{13}C$, $^{15}N$-labeled proteins. J. Biomol. NMR *4*, 231–240.

Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids (New York, USA: John Wiley).

Yoshida, H. (2001). The ribonuclease $T_1$ family. Methods Enzymol. *341*, 28–41.

Yu, Y., Makhatadze, G.I., Pace, C.N., and Privalov, P.L. (1994). Energetics of ribonuclease $T_1$ structure. Biochemistry *33*, 3312–3319.

*Received February 17, 2003; accepted April 7, 2003*

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. 2004, 52, 4969–4974 4969

# Nondestructive Observation of Bovine Milk by NMR Spectroscopy: Analysis of Existing States of Compounds and Detection of New Compounds

FANGYU HU, KAZUO FURIHATA, MIE ITO-ISHIDA, SHUICHI KAMINOGAWA, AND MASARU TANOKURA*

Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

In this study were successfully observed the one- ($^1$H, $^{13}$C) and two-dimensional ($^1$H−$^{13}$C, $^1$H−$^{15}$N, $^1$H−$^{31}$P) NMR spectra of milk directly without any pretreatment. The signals of each NMR spectrum were assigned, and their existing states were also analyzed. Lactose existed in a free state in milk. The signals due to the butyric acid chain can be assigned among the other fatty acid chains. Monounsaturated fatty acid (oleic acid chains) and polyunsaturated fatty acid chains (linoleic and linolenic acid) were assigned by their characteristic signals. The signals from citrate, *N*-acetylcarbohydrates, and lecithin could be observed directly in the $^1$H−$^{13}$C HSQC NMR spectra; the assignment of their signals was made through the $^1$H−$^{13}$C, $^1$H−$^{15}$N, and $^1$H−$^{31}$P HMBC spectra of extracted milk. Signals from creatine and *N*-acetylcarbohydrates were detected for the first time.

**KEYWORDS:** **$^1$H NMR; $^{13}$C NMR; $^{31}$P NMR; 2D-NMR; milk; saturated fatty acid chain; unsaturated fatty acid chain; creatine; *N*-acetylcarbohydrates; lecithin**

## INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy has been widely applied in organic chemistry and biochemistry to identify organic compounds and to structurally analyze biopolymers (*1*). Recently, NMR spectroscopy has been shown to be a very effective and versatile tool for food scientists because it is nondestructive, selective, and capable of simultaneous detection of compounds in complex mixtures (*2*). NMR spectroscopy has, therefore, been applied to the analysis of oils, juice, drugs, and coffee (*3*−*5*). In the observations that have been conducted, NMR experiments have been made under "no water" conditions. NMR experiments have always required the removal of $H_2O$ and the dissolution in $D_2O$. This may cause the loss of volatile or low molecular weight substances.

Other analytical techniques such as chromatography and mass spectrometry have been successfully applied to milk (*6*). All of these methods require specific extractions from milk, which cause problems such as the loss or change of volatile or sensitive compounds and the denaturation of proteins, as well as measurements that require a long time to complete and are complicated to perform. $^1$H NMR spectroscopy has not been applied directly to milk for two reasons: Milk is overwhelmingly made up of water, which means that the signal from water is so large that it possibly overlaps most of the signals in the $^1$H NMR spectra from other components; milk is a complicated emulsion, and therefore it is difficult to obtain a good, sensitive, NMR spectrum. As a result, although there have been several reports on the NMR measurements of milk (*7*−*12*), almost all of the NMR experiments that have been conducted to date have been made by $^{13}$C and $^{31}$P NMR and required pretreatment of the milk, which consisted of the extraction of triacylglycerols, removal of fat and metal ions, or adjustment of the pH (*7*−*9*, *11*, *12*). The $^1$H NMR spectrum of milk has been obtained with the spin−echo pulse sequence, where 0.2 mM $MnCl_2$ was added to milk and the signals were not assigned (*10*). Recently, with the development of NMR machines and the improvement in measurement techniques, it has become possible to apply $^1$H and $^{13}$C one- (1D) and two-dimensional (2D) NMR spectra directly to milk without any additive or pretreatment.

In this study we tested the potential of NMR spectroscopy as a tool for the analysis of milk and investigated the number of constituents that could be observed in the NMR spectra of milk. Milk was analyzed through 1D and 2D NMR experiments. This is the first report on the observation of the NMR spectra of milk without any pretreatment.

## MATERIALS AND METHODS

**Materials and Sample Preparation.** $D_2O$ (99.7%) was purchased from Shoko Co. Ltd. (Tokyo, Japan) and $CDCl_3$ (99%) from Isotec Inc. (Tokyo, Japan). Ultrahigh-temperature pasteurized, homogenized whole milk was purchased at a local supermarket. For most of the measurements, 0.1 mL of $D_2O$ was added to 0.9 mL of milk, to make it easier to adjust the lock and shims system so that better spectra could be easily obtained. The sample was then placed in a 5 mm NMR tube. The volume of the sample was ~0.65 mL. Whole milk without any

* Corresponding author (telephone +81-3-5841-5165; fax +81-3-5841-8023; e-mail amtanok@mail.ecc.u-tokyo.ac.jp).

10.1021/jf049616o CCC: $27.50 
Published on Web 07/08/2004

additive was also examined by NMR to confirm that $D_2O$ has no effect on the $^1H$ and $^{13}C$ NMR spectra of milk.

Two-dimensional HMBC NMR spectra of extracted milk were measured to assign minor signals of whole milk. Skim milk was obtained by removing the fatty acid layer after centrifugation three times at 6000 rpm for 30 min (*13*). Then 2-fold volumes of ethanol (99.5%) purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan), were added to skim milk to denature proteins, and after the mixture was shaken by vortex for 5 min, it was centrifuged for 30 min at 15000 rpm. Then the supernatant was vacuum-dried and dissolved in $D_2O$ (99.7%). With these treatments, most of the fats and proteins were removed from milk. The extracted milk prepared in this way was used for the 2D HMBC NMR experiments.

**NMR Spectroscopy.** All of the NMR experiments were made at room temperature on a JEOL JNM-α 500 NMR spectrometer. The $^1H$ NMR spectra of whole milk were measured at 500 MHz. The signal of $H_2O$ was suppressed by the presaturation method. The $H_2O$ was used as an external reference, and its chemical shift was 4.65 ppm. The number of data points was 16K; the acquisition time was 3.28 s, the delay time, 2.0 s, and the number of scans, 16. The $^{13}C$ NMR spectra of whole milk were recorded at 125.65 MHz. Dioxane was used as an external reference, and its chemical shift was 67.4 ppm. The number of data points was 16K; an acquisition time of 0.54 s and a delay time of 2.0 s gave a total repetition time of 2.54 s, and 1000 scans were accumulated with proton decoupling.

The $^1H-^{13}C$ PFG-HSQC spectra of whole milk were performed with the phase sensitive mode. The acquisition parameters were as follows: number of data points, 1024 for $^1H$ and 256 for $^{13}C$; spectral width, 8000 Hz ($^1H$) and 20169 Hz ($^{13}C$); digital resolution, 7.81 Hz in F2 ($^1H$) and 78.79 Hz in F1 ($^{13}C$); acquisition time, 0.128 s; delay time, 1.5 s; number of scans, 64; number of dummy scans, 16. For the 2D HMBC NMR experiments, the total measurement time required for the $^1H-^{13}C$ operation was 7.5 h.

The $^1H-^{13}C$ FG-HMBC spectra of extracted milk were acquired with the absolute mode. The acquisition parameters were as follows: number of scans, 64; dummy scans, 16; number of data points, 512 in F2 ($^1H$) and 512 in F1 ($^{13}C$); spectral width, 3016.59 Hz in F2 ($^1H$) and 25113.01 Hz in F1 ($^{13}C$); digital resolution, 5.89 Hz in F2 and 49.05 Hz in F1; acquisition time, 0.1697 s; delay time, 1.8 s; HMBC delay time, 60 ms.

The $^1H-^{31}P$ FG-HMBC spectra of extracted milk were performed with the absolute mode. Potassium phosphate was used as an external reference, and its chemical shift of $^{31}P$ was 0 ppm. The acquisition parameters were as follows: number of scans, 16; number of dummy scans, 2; number of data points, 512 in F2 ($^1H$) and 512 in F1 ($^{31}P$); spectral width, 2514.46 Hz in F2 ($^1H$) and 10000 Hz in F1 ($^{31}P$); digital resolution, 4.91 Hz in F2 and 19.53 Hz in F1; acquisition time, 0.2036 s; delay time, 1.4 s; HMBC delay time, 60 ms.

The $^1H-^{15}N$ FG-HMBC spectra of extracted milk were acquired with the absolute mode. Ammonia was used as an external reference, and its chemical shift of $^{15}N$ was 0 ppm. The acquisition parameters were as follows: number of scans, 256; number of dummy scans, 16; number of data points, 512 in F2 ($^1H$) and 256 in F1 ($^{15}N$); spectral width, 3077.87 Hz in F2 ($^1H$) and 20408.16 Hz in F1 ($^{15}N$); digital resolution, 6.01 Hz in F2 and 79.72 Hz in F1; acquisition time, 0.1663 s; delay time, 1.8 s; HMBC delay time, 60 ms.

**Assignment of NMR Signals.** At first, the NMR spectra of milk were analyzed by referring to the published data of chemical shifts for most compounds of milk (*6*, *14*, *15*). The $^1H$ and $^{13}C$ signals of lactose and fats were assigned in such a way. Although the chemical shifts of $^1H$ and $^{13}C$ signals of whole milk are a little different from published data, due to the different solvent conditions, the correlations in 2D NMR spectra would not change. We, therefore, assigned signals in 1D $^1H$ and $^{13}C$ NMR spectra tentatively and then demonstrated the accuracy of assignments in various kinds of 2D spectra. The signals due to trace compounds could be also detected in the 2D NMR spectra. To assign the signals of trace compounds in whole milk, we used the correlations of signals in 2D HMBC spectra. Finally, the assignments were confirmed by adding authentic compounds to whole milk.

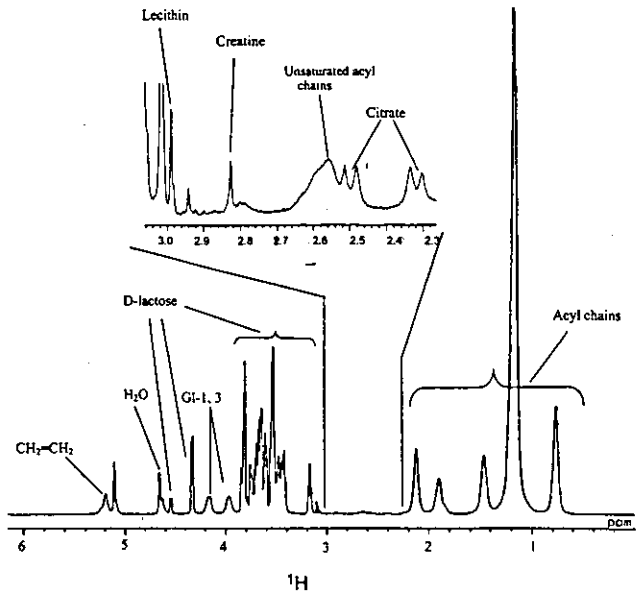


**Figure 1.** $^1H$ NMR spectrum of milk.

## RESULTS AND DISCUSSION

The $^1H$ and $^{13}C$ NMR spectra of milk were obtained both in the absence and in the presence of 10% $D_2O$, which indicated no effect on the NMR spectra from the addition of $D_2O$ (data not shown). However, the addition of 10% $D_2O$ to the milk made it easier to adjust the lock and shims system during the NMR experiments. As a result, a small amount (10%) of $D_2O$ was added to the milk for the NMR measurements. The $^1H$ and $^{13}C$ NMR spectra are shown in **Figures 1** and **2**, respectively.

**$^1H$ NMR Spectrum of Whole Milk.** The $^1H$ NMR spectrum of milk is shown in **Figure 1**. Note that the signal caused by $H_2O$ at 4.65 ppm was successfully suppressed. Three main regions of the $^1H$ NMR spectrum were characterized. A high-field (low-frequency) region between 0.6 and 2.2 ppm was assigned to the signals due to acyl chains of milk fats, and a mid-low-field region between 3.1 and 5.1 ppm was assigned to the signals due to lactose, on the basis of the contents and their chemical shifts. Attention should also be paid to the region of weak signals between 2.2 and 3.1 ppm, because the signals at this region are due to trace compounds of milk. This region was enlarged as the subspectrum in **Figure 1**. The $^1H$ signals of glycerol were observed at 3.98 and 4.18 ppm, and those of the double bond due to unsaturated milk fats were at 5.20 ppm, as assigned by comparing the chemical shifts of the observed signals with those available in the references. The assignment of the $^1H$ NMR spectrum is summarized in **Table 1**.

Very small signals were observed in the region from 5.5 to 9.0 ppm, which may be due to the amide protons of proteins. These signals were very weak and heavily overlapped, which could be considered to be due to the following reasons: (1) The molecular weights of proteins are so large that their molar concentrations are rather low. (2) Caseins are the main milk proteins, and their signals would be very broad and extremely low in peak height because they may exist as large complexes in milk. (3) In the $^1H$ NMR spectrum of whole milk, the signal of $H_2O$ is suppressed with the presaturation pulse, which would decrease the resonance intensity of the broad signals dramatically by spin diffusion while leaving unaffected the sharp signals (*16*).

In the $^1H$ NMR spectrum, the signals due to lactose were observed to be narrow and sensitive, whereas the signals due to milk fats were broad, because lactose is readily soluble in milk, but milk fats consist of various acyl chains and are

**Table 1.** Assignment of $^1H$ and $^{13}C$ NMR Signals of Compounds in Milk

| compound | chemical shift ($\delta$) $^1H$ | chemical shift ($\delta$) $^{13}C$ | assignment[a] |
|---|---|---|---|
| D-lactose | 3.69–3.82 | 60.83 | $CH_2OH$-α6 |
| | 3.69–3.83 | 60.94 | $CH_2OH$-β6 |
| | 3.64 | 61.92 | $CH_2OH$-6′ |
| | 3.80 | 69.44 | CH-4′ |
| | 3.81 | 70.90 | CH-2 |
| | 3.43 | 71.84 | CH-2′ |
| | 3.52 | 72.06 | CH-3 |
| | 3.71 | 72.28 | CH-5 |
| | 3.53 | 73.43 | CH-3′ |
| | 3.15 | 74.74 | CH-2 |
| | 3.51 | 75.22 | CH-3 |
| | 3.49 | 75.58 | CH-5 |
| | 3.60 | 76.15 | CH-5′ |
| | 3.53 | 79.06 | CH-β4 |
| | 3.53 | 79.20 | CH-α4 |
| | 5.03 | 92.68 | CH-α P |
| | 4.48 | 96.61 | CH-β P |
| | 4.27 | 103.67 | CH-1′ |
| acyl chains of fatty acids | 0.78 | 14.63 | $CH_3$-ω1 |
| | 1.19 | 23.34 | $CH_2$-ω2 |
| | 1.47 | 25.45 | $CH_2$-Δ3 |
| | 1.02–1.38 | 29.81–30.50 | $CH_2$ (ω4–n) |
| | 1.02–1.38 | 32.64 | $CH_2$-ω3 |
| | 2.17 | 34.40 | $CH_2$-Δ2 |
| | | 172.15 | $COO^-$ (Gl-2) |
| | | 172.43 | $COO^-$ (Gl-1,3) |
| | 0.78 | 14.46 | butyric $CH_3$-ω1 |
| | 1.47 | 18.81 | butyric $CH_2$-ω2 |
| | 2.17 | 36.10 | butyric $CH_2$-ω3 |
| | 1.89 | 27.78 | oleate $CH_2$-ω8,11 |
| | 5.20 | 130.20 | oleate HC=CH-ω9,10 |
| | 2.68 | 25.52 | linoleate $CH_2$-ω8 |
| glycerol backbone of fats | 3.98, 4.18 | 62.52 | Gl-1, 3 |
| | 5.12 | 69.64 | Gl-2 |

[a] The symbol "ω" indicates the position from the methyl group end; the symbol "Δ" indicates the position from the ester group end.

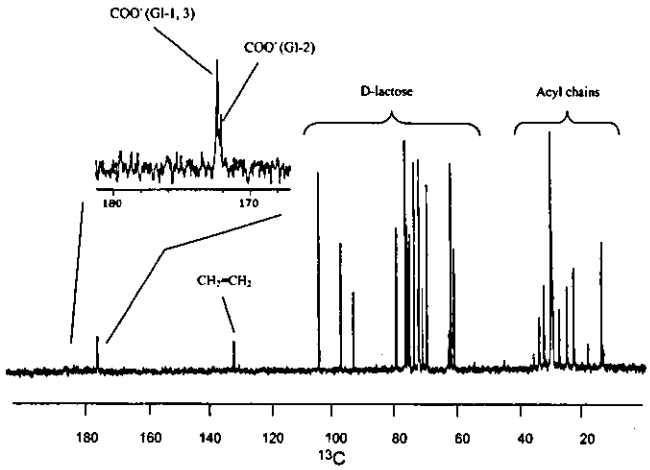


**Figure 2.** $^{13}C$ NMR spectrum of milk.

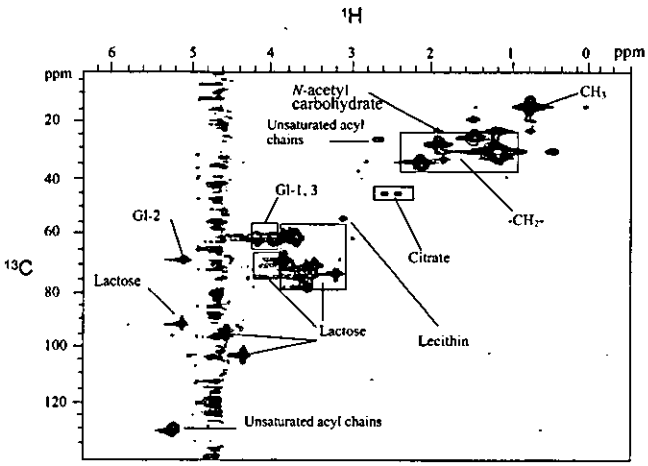


**Figure 3.** $^1H$–$^{13}C$ HSQC NMR spectrum of milk.

suspended in milk as fat globules. We also analyzed the $^1H$ NMR spectra of both nonhomogenized and homogenized milks. Compared with the spectrum of the homogenized milk, the signals caused by the nonhomogenized milk were broader and less sensitive (data not shown), because the nonhomogenized milk contained very large fat globules. As a result, the broadening of the signals and the sensitive level of the signals reflected the size of the fat globules in the milk. However, the $^1H$ NMR spectrum of the milk was very crowded, and many signals overlapped (for example, in the regions containing acyl chains and lactose). Furthermore, the signals between 2.2 and 3.1 ppm were so weak that they could not be assigned by the $^1H$ NMR spectra alone. For further assignment of the overlapped and weak signals, therefore, we measured the $^{13}C$ and $^1H$–$^{13}C$ HSQC spectra.

**$^{13}C$ NMR Spectrum of Whole Milk: Assignment of the Signals of Lactose and Fats.** The $^{13}C$ NMR spectrum of milk is shown in **Figure 2**. The signals were narrow and did not overlap, which allowed us to analyze the spectrum. The details of the signal assignments are summarized in **Table 1**. The signals due to proteins were not observed in the $^{13}C$ NMR spectrum for the same reasons as for the $^1H$ NMR spectrum.

Lactose was assigned after being compared with the published data on chemical shifts (*14*). The assignment was confirmed by the $^1H$ NMR (**Figure 1**) and $^1H$–$^{13}C$ HSQC spectra (**Figure 3**) of milk. Lactose was concluded to exist in the free state in milk, because its signals in milk are not different from the standard data of the authentic lactose solution (*14*).

There are many kinds of acyl chains in triacylglycerols. The signals were assigned by the comparison of the $^{13}C$ NMR spectrum of milk with the published data (*14*, *15*). The signals of the ω1, ω2, and ω3 carbons of fatty acid acyl chains were observed to overlap at 14.63, 23.34, and 32.64 ppm, respectively, whereas the ω1, ω2, and ω3 carbons of the butyric group were observed at 14.46, 18.81, and 36.10 ppm, respectively. Among short-chain fatty acids, butyric acid is an important flavor component of milk. The $^{13}C$ signals of the other methylene groups from ω4 to ω*n* neighboring the carbonyl end or double bond were observed at 29.81−30.50 ppm. The $^{13}C$ signals at 25.45 and 34.40 ppm were, respectively, assigned to Δ3 and Δ2 of the fatty acid acyl chains. From these signals, it was not possible to distinguish whether the acyl chains came from a saturated or unsaturated fatty acid, except for the butyric group in the triacylglycerols, and the number of carbons could also not be identified. The $^{13}C$ signals at 130.20 and 27.78 ppm were assigned to the double-bond carbons and the methylene carbons connecting the double bonds, respectively. Because the largest quantity of unsaturated acids in milk is supplied by oleic acid (*6*), it has been suggested that these two signals are mainly contributed by oleic acid. The signal at 172.15 ppm is caused by the carbonyl group bonding to the 2-position of glycerol, and the signal at 172.43 ppm is caused by the carbonyl group bonding to the 1,3-positions of glycerol (*5*). The signals due to the carbonyl group were enlarged in the subspectrum of the $^{13}C$ NMR spectrum in **Figure 2**.

The signal at 62.52 ppm was assigned to the $C_1$ and $C_3$ carbons of the glycerol moiety of fats. The signal at 69.64 ppm

was assigned to the $C_2$ carbon of the glycerol moiety, although it overlapped with the lactose signals.

Compared with the $^{13}C$ NMR spectrum of triacylglycerols extracted from milk (*11*, *12*), our $^{13}C$ NMR spectrum of whole milk appeared to be somewhat different because fats exist in milk as the fat globules, whereas extracted fats exist in the free state in organic solvent.

**$^1H-^{13}C$ HSQC Spectrum of Whole Milk. Figure 3** shows the $^1H-^{13}C$ HSQC spectrum of whole milk without extraction. The noise from $H_2O$ was not completely suppressed, despite presaturation. The $^1H-^{13}C$ HSQC spectrum supplied useful information on the signals, due to minor constituents that overlapped in the 1D $^1H$ NMR spectra.

The $^1H$ signals at 2.2−3.1 ppm showed correlations to the $^{13}C$ signals at 25.0−55.0 ppm. Through consultation of the databases (*14*, *15*), the $^1H$ resonance at 2.68 ppm, which showed a cross-peak with the $^{13}C$ resonance at 25.52 ppm, was assigned to be the methylene group next to the double bond of polyunsaturated acyl chains. This signal was concluded to be caused by the linoleate and linolenate chains, because they are the main polyunsaturated fatty acids in milk. The $^1H$ resonances at 2.40 and 2.54 ppm are suggested to be due to 2,4-$CH_2$ of citric acid, because they were correlated to the $^{13}C$ signal at 45.15 ppm in the HSQC spectrum. The $^1H$ signals at 1.95 and 3.10 ppm showed correlations to the $^{13}C$ signals at 23.1 and 54.90 ppm, respectively. The signals supplied too little information to be assigned, so we measured the HMBC spectra of extracted milk.

**$^1H-^{13}C$ HMBC Spectrum of Extracted Milk.** HMBC spectroscopy is particularly useful because it connects protons with carbons via two or three bond couplings that can supply more information about connectivities. We attempted to employ HMBC experiments of market milk without any pretreatment at first, but the HMBC spectrum could not be obtained. Extracted milk was, therefore, prepared by removing water and fat.

*Detection of Citrate.* The $^1H-^{13}C$ HMBC spectrum of extracted milk is shown in **Figure 4A**. In the spectrum, the $^1H$ signals at 2.40 and 2.54 ppm showed a correlation with the $^{13}C$ signal at 45.15 ppm, which was also the case in the $^1H-^{13}C$ HSQC spectrum (**Figure 3**). Both of these protons also connected to the $^{13}C$ signals at 77.35, 179.97, and 182.50 ppm in the $^1H-^{13}C$ HMBC spectrum (**Figure 4A**). These data suggest that they are from ionized citric acid, as judged by the chemical structure of citric acid (**Figure 4B**) (*15*). Because the pH of milk was ~6.8, the ionic state of citrate could be predicted from its $pK_a$, and the chemical shifts of signals coincide with those reported for citrate ion. Citrate is the main organic acid in milk and provides the weak acidic flavor, as well as its heat stability in milk (*6*). It has been reported that citrate binds to calcium ion and acts as a component of casein micelles (*6*). However, from the NMR spectra, the line widths of the signals due to the citrate ion in the milk were narrow and showed little difference from those of citrate dissolved in water. Thus, we think that in milk, citrate is mobile to a considerable degree. The results of the assignment of the signals from citrate ion are summarized in **Table 2**.

*Detection of Creatine.* In the $^1H-^{13}C$ HMBC spectrum of extracted milk, a new $^1H$ signal at 2.88 ppm was observed (**Figure 4A**). This $^1H$ resonance was correlated to the $^{13}C$ signal at 55 and 158 ppm in the $^1H-^{13}C$ HMBC spectrum and to the $^{15}N$ signal at 78 ppm in the $^1H-^{15}N$ HMBC spectrum (**Figure 5A**). The other $^1H$ signal at 3.79 ppm was correlated to the $^{13}C$ signals at 38, 158, and 176 ppm in the $^1H-^{13}C$ HMBC spectrum.

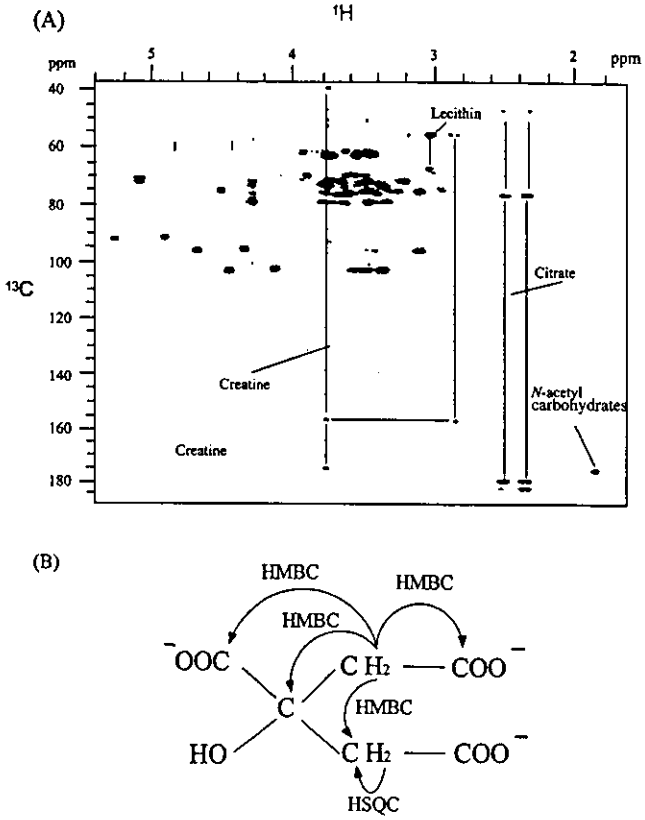


**Figure 4.** (A) $^1H-^{13}C$ HMBC NMR spectrum of the extracted milk. (B) Chemical structure and the expected NMR correlation of citrate. The cross-peaks can be observed among $CH_2$ with C and $COO^-$ in the $^1H-^{13}C$ HMBC spectrum and between C and H of $CH_2$ in the $^1H-^{13}C$ HSQC spectrum.

**Table 2.** Assignment of NMR Signals of Citrate, Creatine, *N*-Acetylcarbohydrates, and Lecithin in Milk

| | chemical shift ($\delta$) | | | | |
|---|---|---|---|---|---|
| compound | $^1H$ | $^{13}C$ | $^{15}N$ | $^{31}P$ | assignment |
| citrate | 2.40, 2.54 | 45.15 | | | $CH_2$ |
| | 2.40, 2.54 | 77.35 | | | C |
| | 2.40, 2.54 | 179.97 | | | $CH_2$–$COO^-$ |
| | 2.40, 2.54 | 182.50 | | | C–$COO^-$ |
| creatine | 2.88 | 55.00 | | | $CH_2$ |
| | 2.88 | 158.00 | | | C |
| | 2.88 | | 78.00 | | N* |
| | 3.79 | 38.00 | | | $CH_3$ |
| | 3.79 | 158.00 | | | C |
| | 3.79 | 176.00 | | | COOH |
| *N*-acetylcarbohydrates | 1.95 | 23.1 | | | $CH_3$ |
| | 1.95 | 176.00 | 120 | | $NHCOCH_3$ |
| lecithin | 3.12 | 54.90 | | | 3Me |
| | 3.12 | 67.50 | | | $CH_2$-5 |
| | 4.22 | | 46.51 | | $CH_2$-4 |
| | 3.12 | | 46.51 | | 3Me |
| | 4.22 | | | 0.6 | $CH_2$-4 |
| | 4.18 | | | 0.6 | $CH_2$-3 |
| | 3.75, 3.83 | | | 0.6 | $CH_2$-5 |

The cross-peaks in the 2D NMR spectra and the chemical shift values of the signals indicated that these signals were caused by creatine from consideration of its chemical structure (**Figure 5B**). The chemical shift values and the narrow line widths of the observed signals compared with the standard data (*15*) led us to the conclusion that creatine exists in a free state, without any interaction with other compounds in milk. The assignments that were obtained are listed in **Table 2**.

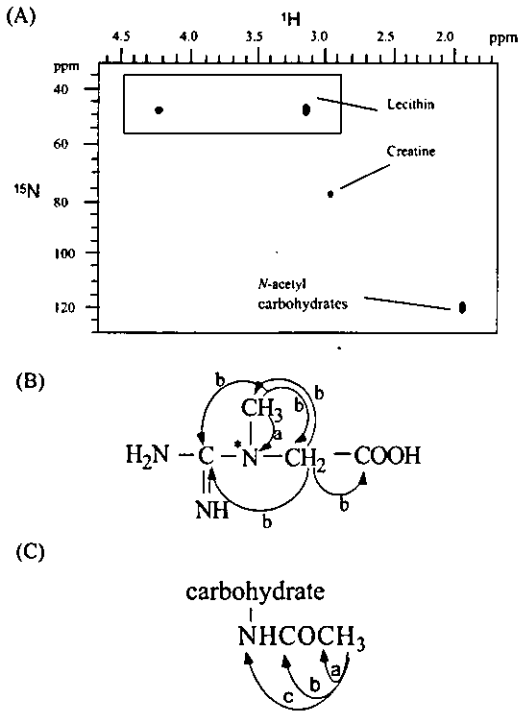


**Figure 5.** (**A**) $^1$H–$^{15}$N HMBC NMR spectrum of the extracted milk. (**B**) Creatine correlations; "a" and "b" indicate the possible correlations in the $^1$H–$^{15}$N HMBC and $^1$H–$^{13}$C HMBC spectra, respectively. (**C**) *N*-Acetylcarbohydrate correlations; "a", "b", and "c" indicate the possible correlations in the$^1$H–$^{13}$C HSQC, $^1$H–$^{13}$C HMBC, and $^1$H–$^{15}$N HMBC spectra, respectively.

*Detection of N-Acetylcarbohydrates.* In the $^1$H−$^{13}$C HSQC spectrum of milk (**Figure 3**), the $^1$H signal at 1.95 ppm, which was overlapped with the signals from fatty acid acyl chains in the 1D $^1$H NMR spectrum, showed a correlation to the $^{13}$C signal at 23.1 ppm. This $^1$H signal was also correlated to the $^{13}$C signal at 176 ppm in the $^1$H−$^{13}$C HMBC spectrum (**Figure 4A**) and correlated to the $^{15}$N signal at 120 ppm in the $^1$H−$^{15}$N HMBC spectrum of the extracted milk (**Figure 5A**). After consultation of some databases (*15*, *17*), it was concluded that these signals were due to the acetamido group of *N*-acetylcarbohydrates (**Figure 5C**). To confirm this, a small amount of *N*-acetylcarbohydrate (*N*-acetyllactosamine or *N*-acetylglucosamine) was added to the sample of extracted milk, and then its $^1$H−$^{13}$C HMBC spectrum was observed. In the spectrum, the cross-peak between 1.95 ppm in $^1$H and 176 ppm in $^{13}$C became larger, but no new signal appeared. Thus, it was confirmed that *N*-acetylcarbohydrates are contained in milk. *N*-Acetylcarbohydrates are very important because they stimulate the production of bifidus bacteria, which enhances intestinal function (*18*). To our knowledge, this is the first time *N*-acetylcarbohydrates have been detected in commercial milk. The assignment is summarized in **Table 2**.

*Detection of Lecithin.* The $^1$H signal at 3.10 ppm showed a correlation to the $^{13}$C signal at 54.90 ppm in the HSQC spectrum of milk and also showed a correlation to the same $^{13}$C signal in the HMBC spectrum of the extracted milk. This proton was also correlated to the $^{13}$C signal at 67.50 ppm in the $^1$H−$^{13}$C HMBC spectrum. These observations supposed that the signals were due to the trimethylamine group (*15*). This hypothesis was confirmed by the $^1$H−$^{15}$N HMBC spectrum (**Figure 5A**), which showed that the $^1$H signal at 3.12 ppm due to the methyl group and that at 4.22 ppm due to the methylene group were coupled to the $^{15}$N signal at 46.51 ppm. A previous paper stated that

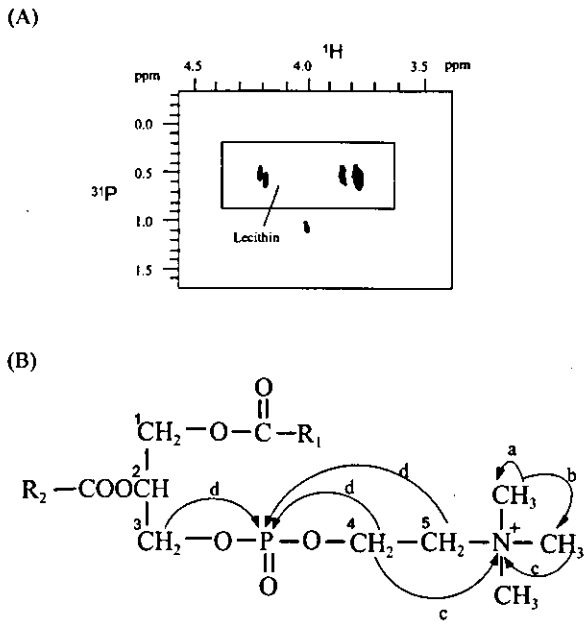


**Figure 6.** (**A**) $^1$H–$^{31}$P HMBC NMR spectrum of the extracted milk. (**B**) Lecithin correlations; "a", "b", "c", and "d" indicate the possible correlations in the$^1$H–$^{13}$C HSQC, $^1$H–$^{13}$C HMBC, $^1$H–$^{15}$N HMBC, and $^1$H–$^{31}$P HMBC spectra, respectively.

trimethylamine was detectable in milk but that the quantity of trimethylamine was very small and even showed undetectable levels in half-numbers of milk (*19*). The signals from trimethylamine could not be observed in the milk we analyzed. In milk, the trimethylamine group mainly exists in lecithin (*6*). The $^1$H−$^{31}$P HMBC spectrum (**Figure 6A**) supported the hypothesis that we observed the trimethylamine group of lecithin, because the proton resonance at 4.22 ppm showed a cross-peak with the $^{31}$P resonance at 0.60 ppm (**Table 2**). The assignment of lecithin was done by detecting the cross-peaks in the $^1$H−$^{13}$C HSQC, $^1$H−$^{13}$C HMBC, $^1$H−$^{15}$N HMBC, and $^1$H−$^{31}$P HMBC spectra, as shown in **Figure 6B**. The signals may be due to glycerophosphocholine and some other similar compounds.

In conclusion, $^1$H, $^{13}$C, and $^1$H−$^{13}$C HSQC NMR spectra of whole milk could be observed successfully without any additive or pretreatment. We confirmed that 10% $D_2O$ added to milk increases the resolution and S/N ratio of NMR signals more easily. The 2D HMBC spectra of extracted milk have been performed to assign signals. During the analysis of milk, we have found the convenience, sensitivity, and efficiency of NMR spectroscopy. Using NMR, we may distinguish the milk of different kinds of cows and other animals. The quantitative analysis of various compounds of milk would also be feasible. This technology is considered to be applicable to quality control or specification of other types of mixtures such as juice, serum, and urine.

## ABBREVIATIONS USED

FG-HMBC, heteronuclear multiple-bond correlation with field gradient; FID, free induction decay; NMR, nuclear magnetic resonance; PFG-HSQC, heteronuclear single-quantum coherence with pulse field gradient; 1D, one-dimensional; 2D, two-dimensional.

## ACKNOWLEDGMENT

We thank Drs. S. Dosako and H. Konishi (of the Snow Brand Milk Products Co., Ltd., Japan) for helpful input.

## LITERATURE CITED


(1) Abraham, R. J.; Fisher, J.; Loftus, P. *Introduction to NMR Spectroscopy*; Wiley: Chichester, U.K., 1988.

(2) Kosir, I. J.; Kidric, J. Identification of amino acids in wines by one- and two-dimensional nuclear magnetic resonance spectroscopy. *J. Agric. Food Chem.* **2001**, *49*, 50−56.

(3) Belton, P. S.; Delgadillo, I.; Gil, A. M.; Casuscelli, F.; Colquhoun, I. J.; Dennis, M. J.; Spraul, M. High field proton NMR studies of apple juices. *Magn. Reson. Chem.* **1997**, *35*, S52−S60.

(4) Bosco, M.; Toffanin, R.; Palo, D.; Zatti, L.; Segre, A. High-resolution $^{1}H$ NMR investigation of coffee. *J. Sci. Food Agric.* **1999**, *79*, 869−878.

(5) Vlahov, G.; Chepkwony, P. K.; Ndalut, P. K. $^{13}C$ NMR characterization of triacylglycerols of moringa oleifera seed oil: An "oleic-vaccenic acid" oil. *J. Agric. Food Chem.* **2002**, *50*, 970−975.

(6) Kaminogawa, S., Sagano, C., Horono, A., Eds. *The Science of Milk—New Function of Milk*; National Plant Dairy Industry of Agricultural Cooperative Association: Tokyo, Japan, 2000 (in Japanese).

(7) Belloque, J.; Carrascosa, A. V.; Fandino, R. L. Changes in phosphoglyceride composition during storage of ultrahigh-temperature milk, as assessed by $^{31}P$-nuclear magnetic resonance: Possible involvement of thermoresistant microbial enzymes. *J. Food Prot.* **2001**, *64*, 850−855.

(8) Ishi, T.; Hiramatsu, K.; Ohba, T.; Tsutsumi, A. The liquid-state $^{31}P$-nuclear magnetic resonance study on microfiltrated milk. *J. Dairy Sci.* **2001**, *84*, 2357−2363.

(9) Belloque, J.; Fuente, M. A.; Ramos, M. Qualitative and quantitative analysis of phosphorylated compounds in milk by means of $^{31}P$-NMR. *J. Dairy Res.* **2000**, *67*, 529−539.

(10) Thomas, M. E.; Robert, G. B. High-resolution proton NMR spectroscopy of milk, organge juice, and apple juice with efficient suppression of the water peak. *J. Agric. Food Chem.* **1986**, *34*, 834−837.

(11) Andreotti, G.; Trivellone, E.; Lamanna, R.; Luccia, A. D.; Motta, A. Milk identification of different species: $^{13}C$-NMR spectroscopy of triacylglycerols from cows and buffaloes' milks. *J. Dairy Sci.* **2000**, *83*, 2432−2437.

(12) Andreotti, G.; Lamanna, R.; Trivellone, E.; Motta, A. $^{13}C$ NMR spectra of TAG: an easy way to distinguish milks from different animal species. *J. Am. Oil Chem. Soc.* **2002**, *79*, 123−127.

(13) Lee, K. H.; Jung, M. Y.; Kim, S. Y. Effects of ascorbic acid on the light-induced riboflavin degradation and color changes in milks. *J. Agric. Food Chem.* **1998**, *46*, 407−410.

(14) Breitmaier, E.; Voelter, W. *Carbon-13 NMR Spectroscopy*; VCH: New York, 1990.

(15) Hayamizu, K.; Yanagisawa, M.; Yamamoto, O. *Integrated Spectral Data Base System for Organic Compounds*; National Institute of Advanced Industrial Science and Technology: Tsukuba, Ibaraki, Japan; accessed March 25, 2004, http://www.aist.go.jp/RIODB/SDBS/.

(16) Akasaka, K.; Konrad, M.; Goody, R. S. Selective spin diffusion. A novel method for studying motional properties of biopolymers in solution. *FEBS Lett.* **1978**, *96* (2), 287−290.

(17) Trehout, D.; Desille, M.; Doan, B. T.; Mahler, S.; Fremond, B.; Certaines, J. D.; Clement, B. Follow-up by one- and two-dimensional NMR of plasma from pigs with ischemia-induced acute liver failure treated with a bioartificial liver. *NMR Biomed.* **2002**, *15*, 393−403.

(18) Nakamura, T.; Amikawa, S.; Harada, T.; Saito, T.; Arai, I.; Urashima, T. Occurrence of an unusual phosphorylated *N*-acetyllactosamine in horse colostrum. *Biochim. Biophys. Acta* **2001**, *1525*, 13−8.

(19) Lunden, A.; Gustafsson, V.; Imhof, M.; Gauch, R.; Bosset, J. O. High trimethylamine concentration in milk from cows on standard diets is expressed as fishy off-flavor. *J. Dairy Res.* **2002**, *69*, 383−390.


**Received for review March 9, 2004. Revised manuscript received June 4, 2004. Accepted June 9, 2004. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan and from the Japan Society for the Promotion of Science.**

JF049616O

*Journal of Biomolecular NMR* **29**: 539–540, 2004.



# Letter to the Editor: $^{1}H$, $^{13}C$ and $^{15}N$ assignments of the tandem WW domains of human MAGI-1/BAP-1

Yusuke Kato[a], Atsushi Akai[a], Rintaro Suzuki[a], Hiroshi Hosokawa[c], Haruaki Ninomiya[c], Tomoh Masaki[c], Koji Nagata[a,b] & Masaru Tanokura[a,*]

[a]*Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences and* [b]*Biotechnology Research Center, University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan;* [c]*Department of Pharmacology, Faculty of Medicine Graduate School of Human and Environmental Studies, Kyoto University, Kyoto 606, Japan*

Received 7 January 2004; Accepted 3 March 2004

*Key words:* β-dystroglycan/MAGI-1/β-catenin complex, Duchenne muscular dystrophy, MAGI-1/BAP-1, NMR assignments, tandem WW domains

## Biological context

WW domains are small modules composed of approximately 30 amino acid residues (Sudol and Hunter, 2000). WW domains contain two highly conserved tryptophan residues and mediate protein-protein interactions by recognizing proline-containing sequences. They are found in cytoskeletal and intracellular signaling proteins such as YAP65, Pin1, dystrophin and Nedd4. WW domains are classified at least into four major groups with regard to their binding specificity. Group I WW domains bind to the PY motif, which contains Pro-Pro-Xaa-Tyr sequence.

MAGI-1 is found in various tissues and is localized in the tight junction of epithelial cells (Ide et al., 1999). MAGI-1, -2, and -3 belong to MAGUK family, which are cytoplasmic scaffold proteins that support the plasma membrane of cells. MAGUKs generally have a few PDZ domains and one SH3 domain whereas MAGIs have five PDZ domains and one or two WW domains. MAGI-1 binds to β-catenin or α-actinin4 through a PDZ domain and to β-dystroglycan through the first WW domain (Patrie et al., 2002; Pirozzi et al., 1997). MAGI-1 binds synaptopodin in glomerular podocytes, which suggests that MAGI-1 may play a role in actin cytoskeleton dynamics within polarized epithelial cells (Patrie et al., 2002).

In the second WW domain of MAGI-1, the second Trp is replaced by Tyr. This Trp is important for the specific recognition of PY motif by Group I WW domains (Kato et al., 2002). However, the second WW domain of MAGI-1 is capable of binding to the PY motif in synaptopodin without the second Trp (Patrie et al., 2002). Conformational analysis would elucidate the underlying mechanism that exploits Tyr residue instead of Trp for the specific recognition of the PY motif in synaptopodin.

*To whom correspondence should be addressed. E-mail: amtanok@mail.ecc.u-tokyo.ac.jp

## Methods and results

The cDNA encoding the tandem WW domains of human MAGI-1 (Ala295 through Ala394) was amplified with PCR and inserted into pET-21b (Novagen). Naturally occurring cysteine residues at 333 and 344 were replaced by serine. *E. coli* BL21(DE3) (Novagen) harboring the above-mentioned plasmid was cultivated at 37 °C in the M9 medium supplemented with 1x basal medium eagle vitamin (Gibco), thiamine and trace elements (Cai et al., 1998). The protein was purified from the soluble fraction of *E. coli* cell lysate using Phenyl Sepharose and Superdex 75 (Amersham). $^{15}N$-labeled and $^{15}N/^{13}C$-labeled proteins were prepared using the modified M9 medium containing 1 g/l $^{15}N$ $NH_4Cl$ and/or 2 g/l $^{13}C$ glucose. The samples for NMR experiments contained 2.5 mM protein in 90% $^{1}H_2O$/10% $^{2}H_2O$ or 100% $^{2}H_2O$ with 10 mM Na-Pi buffer (pH 6.8, not corrected for isotope effects), 100 mM NaCl, 0.05 mM DSS, 0.05% $NaN_3$ and 0.1 mM *p*-ABSF. Microtubes (5-mm outer dia-

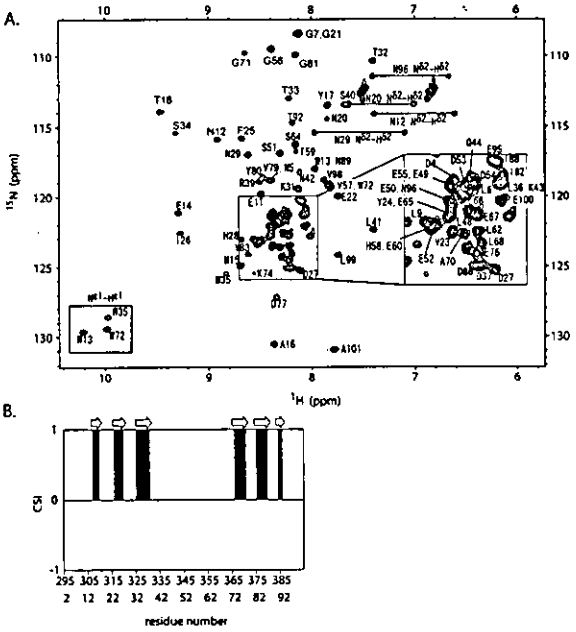


*Figure 1.* (A) Representative strips from the CBCA(CO)NH and HNCACB spectra showing connectivities for residues F318 to W328. (B) Plot of the consensus Chemical Shift Index (CSI) ($^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C'$) of MAGI-1 WW domains, calculated using the program CSI. Indices of −1 and 1 indicate helical structure and β-strand structure, respectively. Expected secondary structural elements are shown as arrows for β-strands.

meter, Shigemi, Tokyo) were used for recording NMR spectra.

All NMR spectra were acquired at 37 °C on a Varian Unity INOVA 500 spectrometer equipped with a Narolac z-axis gradient probe. NMR data were processed and analyzed using NMRPipe (Delaglio et al., 1995) and SPARKY3 (Goddard, T.D. and Kneller, D.G., University of California, San Francisco). $^{1}$H chemical shifts were referenced to internal DSS. $^{13}$C and $^{15}$N chemical shifts were referenced indirectly. The backbone resonance assignments were obtained using $^{1}$H-$^{15}$N HSQC, HN(CO)CA, HNCA, CBCA(CO)NH, HNCACB, HNCO and (HCA)CO(CA)NH experiments. The aliphatic side-chain assignments were obtained using C(CO)NH, H(CCO)NH, HCCH-TOCSY and HCCH-COSY experiments. Secondary structure was predicted using the program CSI (Wishart and Sykes, 1994).

## Extent of assignments and data deposition

Backbone assignment was carried out with the sequential assignment procedure. Signals of backbone amides ($^{15}$N and $^{1}H^{N}$) were unambiguously assigned for 92 residues out of 93 (99%) except for Lys339 in the loop region, which is due to severe overlap (Figure 1A). The tandem WW domain region of MAGI-1 has seven proline residues so that the full-length of our product protein is 101 residue-long including the artificially added N-terminal methionine residue. In addition, resonances of other backbone atoms (98% of $^{13}C^{\alpha}$, 98% of $^{13}C^{\beta}$, 95% of $^{13}C'$ and 90% of $^{1}H^{\alpha}$) as well as those of aliphatic (97% of $^{13}$C, 40% of $^{15}$N and 85% of $^{1}$H) and aromatic (78% of $^{1}$H) side-chain atoms were assigned. We predicted the secondary structure of the tandem WW domains using CSI (Figure 1B). The result indicates that each WW domain contains three β-strands like other WW domains (Kato et al., 2002), while the linker between the WW domains contains no secondary structure. In contrast, tandem WW domains of Prp40 are connected by an α-helical linker, which defines the relative orientation of these domains (Wiesner et al., 2002). Conformational analysis and dynamics study of the tandem WW domains of MAGI-I in the presence and absence of its ligand would show how each WW domain and the linker act in ligand recognition.

The $^{1}$H, $^{13}$C and $^{15}$N chemical shifts have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under accession number 6086.

## Acknowledgements

This work was supported in part by National Project on Protein Structural and Functional Analyses in Japan and Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

Cai, M. et al. (1998) *J. Biomol. NMR*, **11**, 97–102.
Delaglio, F. et al. (1995) *J. Biomol. NMR*, **6**, 277–293.
Ide, N. et al. (1999) *Oncogene*, **18**, 7810–7815.
Kato, Y. et al. (2002) *J. Biol. Chem.*, **277**, 10173–10177.
Patrie, K.M. et al. (2002) *J. Biol. Chem.*, **277**, 30183–30190.
Pirozzi, G. et al. (1997) *J. Biol. Chem.*, **272**, 14611–14616.
Sudol, M. and Hunter, T. (2000) *Cell*, **103**, 1001–1004.
Wiesner, S. et al. (2002) *J. Mol. Biol.*, **324**, 807–822.
Wishart, D.S. and Sykes, B.D. (1994) *J. Biomol. NMR*, **4**, 171–180.