Conformation Control of Peptides by Metal Ions. Coordination Conformation Correlation Observed in a Model for Cys-X-Y-Cys/M^{2+} in Proteins

Takeshi Yamamura,*,† Tsutomu Watanabe,† Atsushi Kikuchi,† Tsutomu Yamane,† Masato Ushiyama,† and Hiroshi Hirota*,‡

Department of Chemistry, Faculty of Science, Science University of Tokyo, Kagurazaka, Shinjuku-ku, Tokyo 162, Japan, and Kawachi Millibioflight Project, ERATO, JST, 4-7-6 Komaba, Meguro-ku, Tokyo 153, Japan

Received August 19, 1996[⊗]

The structure of [(Boc-Cys[1]-Pro-Leu-Cys[4]-OMe)(S-*tert*-C_4H_9)Hg]$^-$ (Boc: butoxycarbonyl), **1**, was studied in *N*,*N*-dimethylformamide (DMF) and compared with that of [(Boc-Cys[1]-Pro-Leu-Cys[4]-OMe)Hg], **2**, in order to discuss the intrinsic structural feature of the cysteine-containing metal-binding sites of proteins: Cys^i-X-Y-Cys^{i+3}/M^{2+}. **1** was generated by the reaction of **2** with NaS-*tert*-C_4H_9. The geometry of the mercury ion (Hg^{2+}) in **1** was proposed to be trigonal planar by UV−vis spectroscopy and Hg L(III) edge X-ray absorption fine structure (XAFS) measurements. Extended X-ray absorption fine structure (EXAFS) calculations yielded r(Hg−S) = 2.42 Å. Analyses of the nuclear Overhauser and exchange spectroscopy (NOESY) and the rotating frame nuclear Overhauser effect spectroscopy (ROESY) spectra of **1** in DMF-d_7 gave approximate distances for the 21 ^{1}H−^{1}H pairs of the main chain loop. These results on distance information were processed by distance geometry (DG) and restrained molecular dynamics (RMD) calculations in order to optimize the molecular structure of **1**. Molecular dynamics (MD) calculations were also performed. We proposed that the trigonal planar Hg^{2+} in **1** regulates the hydrogen-bonding schemes of the peptide in the same manner as the tetrahedral ions involved in the Cys^i-X-Y-Cys^{i+3}/M^{2+} core sites in natural proteins, forming two hydrogen bonds, Cys[1] S−Leu H_N and Cys[1] S−Cys[4] H_N. This is in contrast to **2**, where the linear coordinate mercury causes another type of hydrogen-bonding scheme, Cys[1] S−Leu H_N and Pro CO−Cys[4] H_N. Details of the effect of trigonal planar Hg^{2+} on the peptide conformation were analyzed with respect to the ϕ, φ, and χ torsion angles of the peptide chain. The effect of the change of the ∠S−Hg−S bite angle on the conformation of Cys-Pro-Leu-Cys was also discussed on the basis of MD calculations. The distribution area of Leu (ϕ, φ) in the Ramachandran plot moves from near the α helix region to the turn structure region as the bite angle increases from 90 to 180°, accompanying the change in the hydrogen-bonding scheme. The critical bite angle is around 140°. The analysis revealed that ∠S−Hg−S ≅ 110°, which corresponds to the tetrahedral coordination geometry of the central metal ion, allows a high flexibility of the Cys-Pro-Leu-Cys skeleton.

Introduction

Cys-X-Y-Cys is an amino acid sequence that is widely observed in the inorganic sites of metalloproteins[1] involving Zn^{2+},[2,3] Fe^{2+},[4] Ni^{2+},[5] or Cd^{2+} ions[6] and iron−sulfur clusters.[7,8] This sequence is also found in the functional centers of thioredoxins[9] and glutaredoxins,[10] where it undergoes a reversible transformation between disulfide and sulfhydryl forms, triggering structural changes in neighboring areas. Several conformational studies using model peptides were carried out by Balaram's group[11] with respect to the disulfide forms of the local sequences of these oxidoreductases. Falcomer et al.[12] investigated the structures of Cys-X-Y-Cys model peptides and their redox behavior in order to elucidate the intrinsic conformational tendency of the sequence in proteins. In contrast, the structural chemistry of inorganic models including the Cys-X-Y-Cys sequence have been studied insufficiently.[13] However, it is possible to carry out detailed structural analysis of Cys-X-Y-Cys/metal ion (M^{2+}) systems by a combination of nuclear Overhauser and exchange spectroscopy (NOESY), extended X-ray absorption fine structure (EXAFS) analysis, and molecular dynamics (MD) calculations as demonstrated for zinc finger (ZF) motifs.[3j−o] The chemistry of Cys-X-Y-Cys/M^{2+} accompanying structural analyses allows us to comprehend the correlation between coordination and conformation in metallopeptides and the metal-binding sites in proteins, as well as the intrinsic nature of Cys^i-X-Y-Cys^{i+3}/M^{2+}. It also provides basic knowledge required for the de novo design of artificial metalloproteins.[14]

† Science University of Tokyo.
‡ ERATO, Research Development Corp. of Japan.
⊗ Abstract published in *Advance ACS Abstracts,* September 15, 1997.

(1) (a) O'Halloran, T. V. *Science* **1993**, *261*, 715. (b) Williams, R. J. P. *Biochem. Soc. Trans.* **1990**, *18*, 689.
(2) (a) Gouaux, J. E.; Stevens, R. C.; Lipscomb, W. N. *Biochemistry* **1990**, *29*, 7702. (b) Eklund, H.; Nordstrom, B.; Zeppenzauer, E.; Soderlund, G.; Ohlsson, I.; Borwe, T.; Soderberg, B-O.; Tapia, O.; Branden, C-I.; Akeson, A. *J. Mol. Biol.* **1976**, *102*, 27. Li, H.; Hallows, W. A.; Punzi, J. S.; Marquez, V. E.; Carrell, H. L.; Pankiewitcz, K. W.; Watanabe, K. A.; Goldstein, B. M. *Biochemistry* **1994**, *33*, 23.
(3) (a) Luisi, B. F.; Xu, W. X.; Otwinowski, Z.; Freedman, L. P.; Yamamoto, K. R.; Sigler, P. B. *Nature* **1991**, *352*, 497. (b) Zhang, G.; Kazanietz, M. G.; Blumberg, P. M.; Hurley, J. H. *Cell* **1995**, *81*, 917. (c) Pavletch, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809. (d) Fairall, L.; Schwabe, J. W. R.; Chapman, L.; Finch, J. T.; Rhodes, D. *Nature* **1993**, *366*, 483. (e) Krauris, P. J.; Rain, A. R. C.; Gadhavi, P. L.; Laue, E. D. *Nature* **1992**, *356*, 448. (f) Marmorstein, R.; Carey, M.; Ptashne, M.; Harrison, S. C. *Nature* **1992**, *356*, 408. Baleja, J. D.; Marmorstein, R.; Harrison, S. C.; Wagner, G. *Nature* **1992**, *356*, 450. (g) Berg, J. M. *Annu. Rev. Biophy. Biophys. Chem.* **1990**, *19*, 405. (h) Berg, J. M.; Harrison, S. C. *Nature* **1991**, *353*, 715. (i) Krizek, B. A.; Amann, B. T.; Kilfoil, V. J.; Merkle, D. L.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4518. (j) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635. (k) Härd, T.; Kellenbach, E.; Boelens, R.; Maler, B. A.; Dahlman, K.; Freedman, L. P.; Carlstedt-Duke, J.; Yamamoto, K.; Gustafsson, J-A.; Kaptein, R. *Science* **1990**, *249*, 157. (l) Summers, M. F.; South, T. L.; Kim, B.; Hare, D. R. *Biochemistry* **1990**, *29*, 329. (m) Summers, M. *J. Cell. Biochem.* **1991**, *45*, 41. (n) Palmer, A. G., III; Case, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 9059. (o) Jasanoff, A.; Weiss, M. A. *Biochemistry* **1993**, *32*, 1423.

S0020-1669(96)01009-9 CCC: $14.00

We selected the mercury ion (Hg^{2+}) after preliminary experiments on several metal ions. Among the metal ions studied, Hg^{2+} was superior to the others for several reasons related to our purpose: (1) Hg^{2+} gives diamagnetic compounds in contrast to the biologically important transition metal ions, such as Fe^{2+} [4] and Ni^{2+},[5] and iron−sulfur clusters.[7,8] This enables us to obtain well-defined NMR data. (2) Hg^{2+} is also superior to Zn^{2+} due to its affinity for sulfur ligands. Hg^{2+} forms stable HgS species and has three types of discrete coordination geometry, linear, trigonal planar, and tetrahedral, which are mutually interchangeable in solution.[15] In addition, Hg^{2+} causes no line broadening on the NMR spectra of the Cys-containing ligand which we used in this experiment. This is different from Zn^{2+}, which often causes line broadening on the same ligand.[16] (3) On the other hand, Cd^{2+} also forms stable bonds with Cys-X-Y-Cys; however, it only generates a tetrahedral geometry and yields two isomers.[17] These merits outweighed the drawback that Hg^{2+} is not observed in biological Cys^{i}-X-Y-Cys^{i+3}/M^{2+}, although Hg^{2+} does take an important role in the MerR−Hg^{2+} complex.[15f,18]

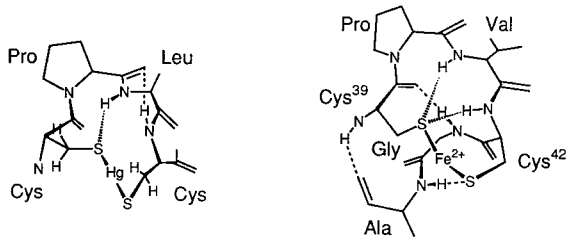

Figure 1. Schematic drawings showing the correlation between coordination and conformation in Cys^{i}-X-Y-Cys^{i+3}/M^{2+} systems: (a) [(Boc-Cys-Pro-Leu-Cys-OMe)Hg], **2**;[19] (b) [(Cys^{i}-Pro-Val-Cys^{i+3})M] in ZFs[3] and Rub.[4]

Previously, we found that the linearly coordinated Hg^{2+} in [(Boc-Cys-Pro-Leu-Cys-OMe)Hg], **2**,[19] realizes a new type of turn structure (Figure 1a) which differs from those of the Cys^{i}-X-Y-Cys^{i+3}/M^{2+} (M = Zn, Fe) of the aspartate transcarbamoylases (AT),[2] zinc finger (ZF) motifs,[3] and rubredoxin (Rub)[4] core sites (Figure 1b). Subsequently, we found that **2** changes its coordination geometry in the presence of *tert*-butyl mercaptan ([S-*tert*-C_4H_9]$^-$) in *N*,*N*-dimethylformamide (DMF). In this paper, we report the results of the Hg L(III) edge EXAFS and the NMR experiments (including NOESY) for the mixture of **2** and [S-*tert*-C_4H_9]$^-$, especially for the 1:1 mixture in DMF-d_7, in which [(Boc-Cys-Pro-Leu-Cys-OMe)(S-*tert*-C_4H_9)Hg]$^-$, **1**, predominates, and those of the structural analysis of **1** using distance geometry (DG), restrained molecular dynamics (RMD) calculations, and MD calculations. We compared the structure of **1** with that of **2** and discussed the effect of the ∠S−Hg−S bite angle on conformation using MD calculations. This work provides the first example of how a coordination change at the

(4) (a) Watenpaugh, K. D.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1979**, *131*, 509. (b) Frey, M.; Sieker, L. C.; Payan, F.; Haser, R.; Bruschi, M.; Pepe, G.; LeGall, J. *J. Mol. Biol.* **1987**, *197*, 525. (c) Adman, E. T.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1991**, *217*, 337.
(5) Volbeda, A.; Charon, M.-H.; Piras, C.; Hatchikian, E. C.; Frey, M.; Fontecilla-Camps, J. C. *Nature* **1995**, *373*, 580.
(6) (a) Vasäk, M.; Kägi, J. H. R. In *Metal Ions in Biological Systems*; Siegel, H., Ed.; Marcel-Dekker: New York, 1983; Vol. 15, p 213. (b) Robbins, A. H.; McRee, D. E.; Williamson, M.; Collet, S. A.; Xoung, N. H.; Furey, W. F.; Wang, B. C.; Stout, C. D. *J. Mol. Biol.* **1991**, *221*, 1269.
(7) (a) Yasunobu, K. T.; Tanaka, M. In *Iron Sulfur Proteins*; Lovenberg, W., Ed.; Academic Press, Inc.: London, 1973; Vol. 2, Chapter 2, p 27. (b) Carter, C. W., Jr. In *Iron Sulfur Proteins*; Lovenberg, W., Ed.; Academic Press, Inc.: London, 1977; Vol. 3, Chapter 6, p 157. (c) Stout, C. D. In *Iron Sulfur Proteins*; Spiro, T. G., Ed.; Wiley-Interscience: New York, 1982; Vol. 4, Chapter 3, p 97. (d) Matsubara, H.; Saeki, K. In *Iron Sulfur Proteins*; Cammack, R., Ed.; Academic Press, Inc.: San Diego, CA, 1992; p 223.
(8) (a) Adman, E. T.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1976**, *251*, 3801. (b) Fukuyama, K.; Matsubara, H.; Tsukihara, T.; Katsube, Y. *J. Mol. Biol.* **1989**, *210*, 383. (c) Barber, M. J.; Neame, P. J.; Lim, L. W.; White, S.; Mathews, F. S. *J. Biol. Chem.* **1992**, *267*, 6611. (d) Stout, C. D. *J. Mol. Biol.* **1989**, *205*, 545.
(9) (a) Dyson, H. J.; Holmgren, A.; Wright, P. E. *Biochemistry* **1989**, *28*, 7074. (b) Katti, S. K.; LeMaster, D. M.; Eklund, H. *J. Mol. Biol.* **1990**, *212*, 167. (c) Dyson, H. J.; Gippert, G. P.; Case, D. A.; Holmgren, A.; Wright, P. E. *Biochemistry* **1990**, *29*, 4129.
(10) (a) Sodano, P.; Xia, T.-H.; Bushweller, J. H.; Björnberg, O.; Holmgren, A.; Billeter, M.; Würtlich, K. *J. Mol. Biol.* **1991**, *221*, 1311. (b) Sodano, P.; Chary, K. V. R.; Björnberg, O.; Holmgren, A.; Kren, B.; Fuchs, J. A. *Eur. J. Biochem.* **1991**, *200*, 369.
(11) (a) Venkatachalapathi, Y. V.; Prasad, V. V.; Balaram, P. *Biochemistry* **1982**, *21*, 5502. (b) Ravi, A.; Venkataram Prasad, B. V.; Balaram, P. *J. Am. Chem. Soc.* **1983**, *105*, 105. (c) Rao, B. N. N.; Kumar, A.; Balaram, H.; Ravi, A.; Balaram, P. *J. Am. Chem. Soc.* **1983**, *105*, 7423. (d) Kishore, R.; Ishizaki, H.; Tu, A. T.; Ravi, A.; Balaram, P. *Int. J. Peptide Protein Res.* **1987**, *30*, 474. (e) Kishore, R.; Raghothama, S.; Balaram, P. *Biopolymers* **1987**, *26*, 873. (f) Kishore, R.; Raghothama, S.; Balaram, P. *Biochemistry* **1988**, *27*, 2462.
(12) Falcomer, C. M.; Meinwald, Y. C.; Choudhary, I.; Talluri, S.; Milburn, P. J.; Clardy, J.; Scheraga, H. A. *J. Am. Chem. Soc.* **1992**, *114*, 4036.
(13) Only a few cases were reported by Nakamura and Ueyama. They focused their attention on the correlation between the Cys-S−H_N hydrogen bond and the redox behavior of Cys-X-Y-Cys/Fe^{2+} ($[Fe_4S_4]^{2+}$) in terms of rubredoxins (Rub)[4] and ferredoxins (Fd).[7,8] The conformational analyses of their compounds were incomplete because of the paramagnetism of their compounds: (a) Ueyama, N.; Nakata, M.; Fuji, M.; Terakawa, T.; Nakamura, A. *Inorg. Chem.* **1985**, *24*, 2190. (b) Ueyama, N.; Kajiwara, A.; Terakawa, T.; Ueno, S.; Nakamura, A. *Inorg. Chem.* **1985**, *24*, 4700. (c) Ueyama, N.; Sugawara, T.; Tatsumi, K.; Nakamura, A. *Inorg. Chem.* **1987**, *26*, 1978. (d) Sun, W.-Y.; Ueyama, N.; Nakamura, A. *Inorg. Chem.* **1993**, *32*, 1095.
(14) Struthers, M. D.; Cheng, R. P.; Impeliali, M. D. *J. Am. Chem. Soc.* **1966**, *118*, 3073.
(15) (a) Liesk, J.; Klar, G. *Z. Anorg. Allg. Chem.* **1977**, *435*, 103. (b) Wertheim, E. *J. Am. Chem. Soc.* **1929**, *51*, 3661. (c) Bradley, D. C.; Kuncher, N. R. *J. Chem. Phys.* **1964**, *40*, 2258. (d) Lechner, J. *Ber. Deut. Chem. Ges.* **1915**, *48*, 1425. (e) Bowmaker, G. A.; Dance, I. G.; Dobson, B. C.; Rogers, D. A. *Aust. J. Chem.* **1984**, *37*, 1607. (f) Watton, S. P.; Wright, J. G.; MacDonnell, F. M.; Bryson, J. W.; Sabat, M.; O'Halloran, T. *J. Am. Chem. Soc.* **1990**, *112*, 2824. (g) Choudhury, S.; Dance, I. G.; Gurney, P. J.; Rae, A. D. *Inorg. Chim. Acta* **1983**, *70*, 227. (h) Cheesmam, B. V.; Arnold, A. P.; Rabenstein, D. L. *J. Am. Chem. Soc.* **1988**, *110*, 6359.
(16) The ^{1}H NMR fingerprint region was observed for Boc-Cys-Pro-Leu-Cys-OMe/Zn^{2+} in DMF-d_7, DMSO-d_6, acetonitrile-d_3, acetone-d_6, and THF-d_8. The spectra of the compound in the latter three solutions were not defined, giving broad signals and many narrow ones. This result suggests that the solutions involve ambiguous mixtures of many conformers and their interacting forms, which probably increases the viscosity of the solutions and shortens the relaxation times of the bands giving rise to line broadening. In contrast, the signal patterns in DMF-d_7 and DMSO-d_6 were comparatively well defined suggesting that the signals originated from a homogeneous system, and the chemical shifts of the bands in the fingerprint regions were in good agreement with those of the trigonal mercury species, **1**, although the bands were still broadened. The reason for this line broadening in the spectrum of Boc-Cys-Pro-Leu-Cys-OMe/Zn^{2+} is still not known. It seems to depend on solvent basicity, the amount of Zn^{2+}, and some unknown factor. Probably these factors are related to each other, as well as to the rapid ligand exchange in Cys-S/Zn^{2+} systems.
(17) Unpublished data. We found that the Boc-Cys-Pro-Leu-Cys-OMe/Cd^{2+} = 1/1 solution only generates [(Boc-Cys-Pro-Leu-Cys-OMe)$_2$-Cd]. The existence of isomers for [(Cys-X-Y-Cys)$_2$M] was proposed by Nakamura and Ueyama: Nakamura, A.; Ueyama, N. *Adv. Inorg. Chem.* **1989**, *33*, 39.
(18) O'Halloran, T. V.; Frantz, B.; Shin, M. K.; Ralston, D. M.; Wright, J. G. *Cell* **1989**, *56*, 119. Shewchuk, L. M.; Verdine, G. L.; Walsh, C. T. *Biochemistry* **1989**, *28*, 2331. Frantz, B.; O'Halloran, T. V. *Biochemistry* **1990**, *29*, 4747. Heltzel, A.; Lee, I. W.; Totis, P. A.; Summers, A. O. *Biochemistry* **1990**, *29*, 9572. Ansari, A. Z.; Bradner, J. E.; O'Halloran, T. V. *Nature* **1995**, *374*, 371.
(19) (a) Yamamura, T.; Watanabe, T.; Kikuchi, A.; Ushiyama, M.; Kobayashi, T.; Hirota, H. *J. Phys. Chem.* **1995**, *99*, 5525. (b) [(Boc-Cys-Pro-Leu-Cys-OMe)Hg] and [(Boc-Cys-Pro-Leu-Cys-Gly-Ala-OMe)Hg] adopt the hybrid conformation composed of the mirror image of γ-turn and ordinary Cys-X-Y-Cys/M^{2+} (M = Zn, Fe) structures. Yamamura, T.; Arai, M.; Yamane, T.; Ukai, T.; Ushiyama, M.; Hirota, H. *Bull. Chem. Soc. Jpn.* **1996**, *69*, 2221.

inorganic site modifies the peptide conformation in the metallopeptide model with a minimal amino acid sequence.

Experimental Details, Data Processing, and Calculations

Preparation of Compounds. All treatments of [(Boc-Cys-Pro-Leu-Cys-OMe)Hg], **2**,[19] and Na[S-*tert*-C_4H_9] were performed under argon (Ar). The solution samples for NMR and XAFS experiments were prepared by reacting [(Boc-Cys-Pro-Leu-Cys-OMe)Hg], **2**, with controlled amounts of Na[S-*tert*-C_4H_9] in *N*,*N*-dimethylformamide-d_7 (DMF-d_7) and DMF (freshly distilled over CaH_2 and BaO), respectively. The standards for XAFS, $[Hg(SCH_3)_2]$,[15b,c] $[Hg(SC_6H_5)_2]$,[15d] (NEt_4)-$[Hg(S\text{-}tert\text{-}C_4H_9)_3]$,[15e,f] and $(NEt_4)_2[Hg(SC_6H_4Cl)_4]$[15g] were prepared according to the methods in the literature. $HgSO_4 \cdot 2H_2O$ and $HgCl_2$ for XAFS standards were obtained commercially. DMF-d_7 (99.5%) and dimethyl sulfoxide (DMSO-d_6; 99.96%) were purchased from Aldrich Inc.

UV−Vis. UV−vis spectra were recorded on a Hitachi 228A spectrophotometer using 1.0-mm optical path length cells in an Ar atmosphere.

NMR. DMF-d_7 samples for NaS-*tert*-C_4H_9/**2** = 0/1, 0.24/1, 0.47/1, 0.73/1, 0.96/1, 1.45/1, 2/1, 4/1, and 11/1 containing 0.041−0.044 mol/dm^3 of **2** were prepared in an Ar box. DMSO-d_6 samples for NaS-*tert*-C_4H_9/**2** = 0/1 and 1/1 samples were also prepared under Ar. The majority of standard NMR spectra (^{1}H, ^{13}C, ^{1}H−^{1}H COSY, ^{1}H−^{13}C COSY, DEPT) for band assignment were obtained using a JEOL GSX 500 spectrometer. Phase-sensitive (PS) NOESY[20] and rotating frame nuclear Overhauser effect spectroscopy (ROESY)[21] data for NaS-*tert*-C_4H_9/**2** = 1/1 samples in DMF-d_7 were acquired on a Bruker ARX500 spectrometer at 298 K with TPPI.[22,23] The DMSO-d_6 solution sample was also observed by NOESY. NOESY data were collected in a 1024 × 1024 data matrix with 16 transients per t_1 point, a sweep width of 6493.41 Hz, an increment of 77 μs of t_1, a mixing time of 0.6 s, and a waiting time of 3.8 s. A squared sine filter shifted by 90° was used both in the t_1 and t_2 dimensions prior to Fourier transformation. The mixing time was selected so as to give a positive NOE after a preliminary experiment on the DMSO-d_6 sample. ROESY spectra were obtained from a 2048 (t_1) × 1024 (t_2) data matrix. The spectrum was acquired with 16 transients for each t_1 point, an increment of 77 μs for t_1, a mixing time of 0.2 s, and a waiting time of 3.0 s. Transmitter power level for spinlock was set so as to obtain a width of 112.5 μs for the second 90° pulse.

XAFS Data Collection, Reduction, and Analysis. The solution samples of NaS-*tert*-C_4H_9/**2** = 0/1, NaS-*tert*-C_4H_9/**2** = 1/1, and NaS-*tert*-C_4H_9/**2** = 2/1 containing 0.1 mol/dm^3 of **2** were sealed under Ar in 3- and 5-mm optical path length cells with polyethylene windows. The reference samples were sealed in polyethylene pellets (under Ar, in the case of thiolates). All samples were run on the beamline BL-10B at the Photon Factory of the National Institute of High Energy Physics (KEK). Measurements were performed under conditions of 2.5 GeV and 280−380 mA using a Si(311) channel-cut double-crystal monochromator. All the XAFS data on the Hg L(III) edge were collected (after blank tests using samples or gold foil set in front of the entrance slit to check harmonics) in the conventional absorption mode at room temperature at the rate of 3−5 s/point using the EXAFS facility of the beamline with N_2/Ar(85/15)- and N_2/Ar(50/50)-filled ion chambers (17 cm) of standard design, for I_0 and I, respectively, operating in a linear regime. Energies were calibrated using gold foil (instead of some mercury compounds). The scans were inspected and rejected if the signal-to-noise level was too high due to beam instabilities or poor detector statistics.

The EXAFS signals were analyzed according to the literature methods.[24,25] The raw spectra observed in energy space (eV) were reduced to the raw EXAFS function in the photoelectron wave vector (k: up to 16.2 Å^{-1}) using the equation $k = [2m(E - E_0)/h^2]^{1/2}$ (E_0: the threshold energy), where the contributions from the pre-edge regions were approximated using Victoreen's function plus constants,[26] and the atomic-like backgrounds were determined by the cubic-spline method. The Fourier transform (FT) of EXAFS signals ($k^3\chi(k)$: k = 3.8 to 11.6 or 11.7 Å^{-1}) to radial distribution function (RDF) and the back-FT from RDF to EXAFS functions (r = 1.5 or 1.45 Å to 2.5 or 2.6 Å) were carried out using the Hanning window with $^1/_{20}$ of the FT ranges. Fourier filtering techniques were used for the EXAFS and radial distribution functions to remove noise. Both curve fitting and ratio methods[27] based on a single scattering model[28] were applied to the first shells of the Hg using the experimental phase shifts and scattering amplitudes extracted from the analyses of the standards, $[Hg(SCH_3)_2]$,[15b,c] $(NEt_4)[Hg(S\text{-}tert\text{-}C_4H_9)_3]$,[15e,f] and $(NEt_4)_2[Hg(SC_6H_4Cl)_4]$.[15g] Curve fittings were refined for $R = \sum\{k^3\chi_{obs} - k^3\chi_{calc}\}^2/\sum\{k^3\chi_{obs}\}^2$. The ratio method was refined by the least-squares method.

Curve fitting and ratio methods are inadequate for treating the RDFs at higher shells, which includes multiple scattering, and for isolating EXAFS backscatterings of particular sets of atoms from those of other atoms at similar distances. To avoid these defects, to thereby distinguish more explicitly the coordination geometries of the mercury centers of NaS-*tert*-C_4H_9/**2** = 0/1 and NaS-*tert*-C_4H_9/**2** = 1/1, and to obtain the absolute Debye−Waller-like factors (σ) for known compounds on the basis of their molecular structures determined by X-ray crystallography, we performed ab initio calculations of theoretical χ data using the FEFF 6.01 program developed by Rehr and his co-workers.[29,30] The core hole reduction factor $S0^2$ was set at 0.85 according to the literature.[31] Θ_D[32] and the threshold energy were the parameters varied in the refinements. The molecular structures were fixed by referring to the experimental results. The obtained theoretical χ data were treated with the program EXAFSH.[25]

Experimental Results

HgS_2 ↔ HgS_3 Equilibrium. Compound **2** alone shows almost no absorption maximum in the UV region (220−350 nm) in MeOH; however, the solution gives a strong LMCT band at λ_{max} = 240 nm when NaS-*tert*-C_4H_9 is added, and the absorption intensity depends on the amount of thiolate added (Figure 2). The appearance of this band indicates that the coordination environment of the new species is similar to that

(20) Jeener, J.; Meier, B. H.; Bauchmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546.

(21) (a) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811. (b) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *63*, 207. (c) Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1987**, *109*, 607.

(22) Marion, D.; Würthrich, K. *J. Magn. Reson. Commun.* **1983**, *113*, 967.

(23) Heteronuclear multiple quantum coherence correlation (HMQC) and heteronuclear multiple-bond correlation (HMBC) were also examined using a Bruker ARX500 spectrometer to assign carbonyl carbons, as well to obtain the evidence for the existence of Cys^1 S−Leu H_α hydrogen bonding. Müller, L. *J. Am. Chem. Soc.* **1979**, *101*, 4481. Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093.

(24) Teo, B. K. *EXAFS: Principles and Data Analysis*; Springer-Verlag: Berlin, 1986.

(25) Kosugi, N.; Kuroda, H. *EXAFS 2, SRI Rep.* **1985**; No. 2. Yokoyama, T.; Hamamatsu, H.; Ohta, T. Program EXAFSH, version 2.1, The University of Tokyo, 1994.

(26) Victoreen, J. A. *J. Appl. Phys.* **1948**, *19*, 855.

(27) Stern, E. A.; Sayers, D. E.; Lytle, F. W. *Phys. Rev. B.* **1975**, *11*, 4836.

(28) Sayers, D. E.; Stern, E. A.; Lytle, F. W. *Phys. Rev. Lett.* **1971**, *27*, 1204.

(29) Rehr, J. J.; Albers, R. C.; Zabinski, S. I. *Phys. Rev. Lett.* **1992**, *69*, 3397. Rehr, J. J.; Mustre de Leon, J.; Zabinsky, S. I.; Albers, R. C. *J. Am. Chem. Soc.* **1991**, *113*, 5135. FEFF enables ab initio multiple scattering calculations based on a theoretical model and removes the drawbacks of the foregoing theoretical standards in which curved wave corrections and electron self-energy were ignored. O'Day et al. demonstrated the usefullness of FEFF version 5 using several known compounds including first transition metal ions.

(30) O'Day, P. A.; Rehr, J. J.; Zabinsky, S. I.; Brown, G. E., Jr. *J. Am. Chem. Soc.* **1994**, *116*, 2938.

(31) Stern, E. A.; Heald, S. M.; Bunker, B. *Phys. Rev. Lett.* **1979**, *42*, 1372.

(32) As FEFF was developed for the EXAFS calculation of solid materials, the σ derived using the Debye temperature (Θ_D) corresponds to the atomic vibration of the phonon mode. If we substitute the IR vibrational frequency of the Hg−S stretching mode of $[Hg(SCH_3)_2]$ (Iwasaki, N.; Tomooka, J.; Toyoda, K. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 1323) into the equation of Debye temperature ($\Theta_D = h\omega_D/k_B$), we obtain Θ_D = 470 K. It is a matter of course that this substitution has no physical meaning. In practice, σ of compound **2** was not simulated by Θ_D = 470 K but by Θ_D = 594 K.

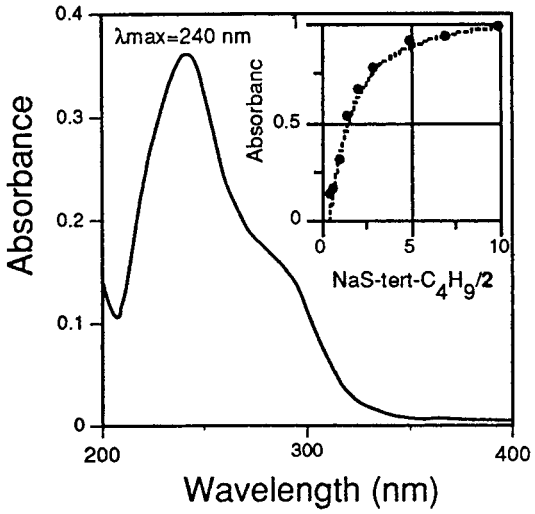

Figure 2. Absorption spectrum of NaS-*tert*-C_4H_9/**2** = 1/1 in MeOH. (The concentration of **2** is 4.9×10^{-4} mol dm^3.) The insert depicts the experimental (●) and the theoretical (dashed line) intensity changes of the absorption maximum (λ_{max} = 240 nm) against the amount of NaS-*tert*-C_4H_9. The dashed line is the best fit obtained using an equilibrium constant of (K) 3.4×10^3 dm^3 mol^{-1} and an extinction coefficient of (ϵ) 1.54×10^4 dm^2 mol^{-1} from eq 1.

of the trigonal planar Hg^{2+} with thiolato ligands and that of the MerR−Hg^{2+} complex.[15f] This result leads to the simple assumption that the reaction represented by eq 1,which involves

$$\mathbf{2} + \text{NaS-}tert\text{-}C_4H_9 \leftrightarrow \text{Na}\mathbf{1} \qquad (1)$$

the equilibrium between bis- and tris(thiolato) mercury(II) complexes,[15h] occurs in the solution. The least-squares fit of the absorption increase on the basis of eq 1 afforded an equilibrium constant (K) of 3.4×10^3 dm^3/mol and an absorption coefficient ($\epsilon_{240\ nm}$) of 1.54×10^4 dm^2/mol (Figure 2).

Figure 3 shows the ^{1}H NMR signal dependence on the NaS-*tert*-C_4H_9/**2** ratio in DMF-d_7 (270 MHz, 298 K). Plots of the chemical shifts (δ) of amide protons (Cys^1 H_N, Leu H_N, Cys^4 H_N) on NaS-*tert*-C_4H_9/**2** clarify that a linear dependence exists in the region NaS-*tert*-C_4H_9/**2** = 0/1 to 1/1. This indicates that an equilibrium such as that represented by eq 1 exists in the DMF-d_7 solution, also, and the equilibrium is shifted to the left as we pointed out from the results of UV−vis spectral measurements in MeOH. Figure 3 also indicates that the proton exchange between **1** and **2** is rapid on the NMR time scale. The linearity is lost in NaS-*tert*-C_4H_9/**2** ≥ 1/1 to 1.5/1. On the other hand, several new bands, which probably originate from plural species, appear in NaS-*tert*-C_4H_9/**2** ≥ 2/1 to 11/1 (not shown here) in addition to the bands observed in the NaS-*tert*-C_4H_9/**2** ≤ 1.5/1 solutions accompanying line broadening. The chemical shifts of amides protons dependent on NaS-*tert*-C_4H_9/**2** were analyzed in the region NaS-*tert*-C_4H_9/**2** ≤ 1/1 by assuming the applicability of eq 1.[33] This analysis afforded $K \geq 5700$ dm^3/mol for Leu H_N in DMF-d_7. This K value is the same order as that derived from the UV−vis experiment for the MeOH solution and indicates that the equilibrium 1 in DMF-d_7 is also shifted strongly to the right as in MeOH. From this K value, the NaS-*tert*-C_4H_9/**2** = 1/1 solution is estimated to contain **1** with the molar ratio of more than 90%.

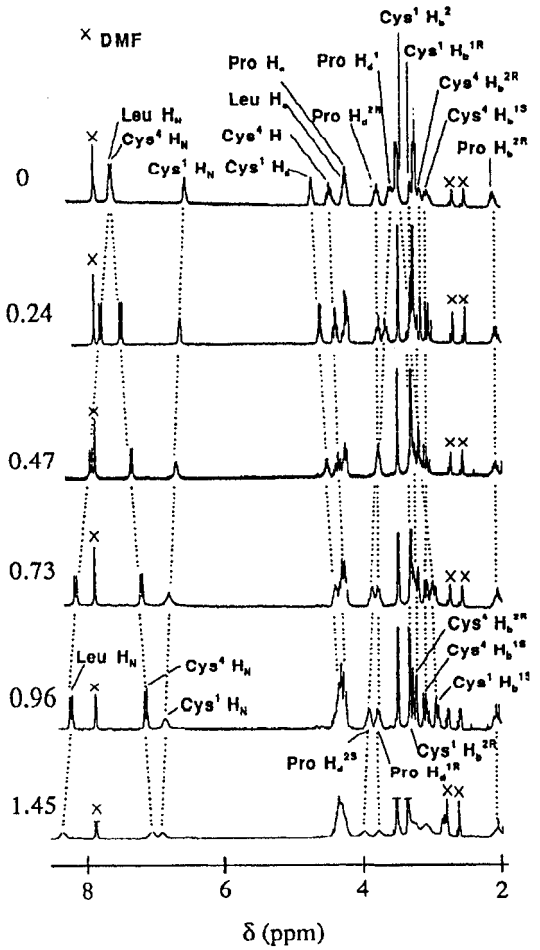

Figure 3. ^{1}H NMR spectra (270 MHz) of NaS-*tert*-C_4H_9/**2** = 0, 0.24, 0.47, 0.73, 0.96, and 1.45 in DMF-d_7. (The concentration of **2** was controlled to 0.044 mol dm^{-3}.)

The low-temperature experiment for the NaS-*tert*-C_4H_9/**2** = 1/2 solution, which was expected to involve both **1** and **2** at a ratio of ≥2/3, failed to eliminate the coagulation between the corresponding amide ^{1}Hs of **1** and **2** even at 223 K, proving the existence of rapid exchange between **1** and **2**.

XAFS. The X-ray absorption near-edge structures (XANES) of the Hg L(III) of NaS-*tert*-C_4H_9/**2** = 0/1, 1/1, and 2/1 samples in DMF were compared with those of the standards (Supporting Information). The XANES spectra of the 1/1 and 2/1 samples clearly differed from that of the 0/1 sample (i.e. **2**) but are close to that of $(NEt_4)[Hg(S\text{-}tert\text{-}C_4H_9)_3]$,[15e,f] suggesting Hg coordination environments in the 1/1 and 2/1 solutions similar to that of $[Hg(S\text{-}tert\text{-}C_4H_9)_3]^-$.

The content of **1** in the 1/1 solution, ≥90% although not 100%, provides a substantial basis for the EXAFS analyses of the solution. The results of the EXAFS analyses for the NaS-*tert*-C_4H_9/**2** = 1/1 solution are summarized in Table 1.[34] The solid lines in Figure 4a,b represent the observed EXAFS and FT of NaS-*tert*-C_4H_9/**2** = 1/1 in DMF, respectively. The corresponding spectra for **2** are shown for comparison in Figure 4c,d by solid lines. Calculations assuming the existence of the Hg−O bond were not optimized rationally. Table 1 indicates

(33) In the case of rapid exchange, the chemical shift ω of the solution which is described by eq 1 is calculated using $[B_0]/(\omega - \omega_2) = 1/K(\omega_1 - \omega) + [A_0]/(\omega_1 - \omega_2)$, where $[B_0]$ and $[A_0]$ are the concentrations of NaS-*tert*-C_4H_9 and **2**, respectively. ω_1 and ω_2 are the chemical shifts of **1** and **2**, respectively. The application of this equation for several NaS-*tert*-C_4H_9/**2** ratios enable us to calculate both K and ω_2. When the chemical shifts of the NaS-*tert*-C_4H_9/**2** = 1.5/1 sample were substituted for ω_2, we obtained $K \cong 300$ dm^3/mol and the content of **1** as ≅ 70%. However, the NaS-*tert*-C_4H_9/**2** = 1.5/1 solution should be excluded from the calculation, as we believe that the solution is affected by a more complicated equilibrium for the NaS-*tert*-C_4H_9/**2** ≥ 2/1 solutions. Therefore, the calculations of K and ω_2 were performed for the solutions with NaS-*tert*-C_4H_9/**2** ≤ 0.75/1, which produced a K larger than 5700 dm^3/mol.

Table 1. Results of the EXAFS Analyses for NaS-*tert*-C_4H_9/**2** = 1/1 in DMF[a]

no.	N[b]	r[c]	$100\Delta C_2$[d]	σ[e]	$1000\Delta C_3$[f]	R[g]	models	methods[h]
1	2.8	2.42	0.2131	0.0750		0.0724	A[i]	CF
2	3.2	2.42	0.2268	0.0767		0.0325	B[j]	CF
3	3.5	2.42	0.0440	0.0820		0.0907	C[k]	CF
4	2.8	2.44	0.2001	0.0742	0.2803	0.0841, 0.0487	A[i]	RA
5	3.1	2.43	0.2053	0.0752	0.0228	0.0858, 0.1799	B[j]	RA
6	3.4	2.41	0.0174	0.0804	−0.3925	0.4799, 0.0163	C[k]	RA
mean	3.2	2.42		0.0773				

[a] Analyses were performed in the FT k range of 3.8 to 11.6 Å^{-1} and back FT r range of 1.5 to 2.5 Å. [b] Number of backscattering atoms. Errors are estimated to be ca. 25% (Cramer, S. P.; Hodgson, K. O.; Stiefel, E. L.; Newton, W. E. *J. Am. Chem. Soc.* **1978**, *100*, 2748). [c] Å. Distance between Hg and S. Errors are estimated to be ca. ± 0.03 Å (see reference in footnote *b*). [d] $\Delta C_2 = (\sigma^2 - \sigma_{mdl}^2)$: σ, Debye−Waller-like factor for unknown compound; σ_{mdl}, Debye−Waller-like factor for model compound, set at 0.00 Å. [e] Debye−Waller-like factor calculated from ΔC_2, obtained here, and σ_{mdl}, obtained by FEFF simulation for models, using the equation $\sigma = \Delta C_2 + \sigma_{mdl}$. [f] ΔC_3: Linear coefficient for the least-squares fitting of phase shift in the ratio method. [g] Refinement parameters: Parameters for curve fitting were defined as $R = \sum\{k^3\chi_{obs} - k^3\chi_{calc}\}^2/\sum\{k^3\chi_{obs}\}^2$. Refinement parameters for the ratio method were those of least-squares fittings for amplitude (the former) and phase shift (the latter). [h] CF, curve fitting; RA, ratio method. [i] $[Hg(SCH_3)_2]$: FT k range, 3.8−11.7 Å^{-1}; back FT r range, 1.5−2.45 Å. [j] $(NEt_4)[Hg(S\text{-}tert\text{-}C_4H_9)_3]$: FT k range, 3.8−11.7 Å^{-1}; back FT r range, 1.6−2.5 Å. [k] $(NEt_4)_2[Hg(SC_6H_4Cl)_4]$: FT k range, 3.8−11.7 Å^{-1}; back FT r range, 1.5−2.6 Å.

that Hg^{2+} in the NaS-*tert*-C_4H_9/**2** = 1/1 solution is bound to three sulfur atoms with Hg−S = 2.42(±0.03) Å and σ = 0.0773 Å.

FEFF calculations were performed using the C_{2h} $Hg(SC)_2$ and C_{3h} $Hg(SC)_3$ cores shown in Figure 5a,b for **2** and **1**, respectively. The bond lengths were taken from the results of the EXAFS analyses in Table 1 (for $Hg(SC)_3$) and our previous study on **2**[19] (for $Hg(SC)_2$). The angles were determined with reference to the results of MNDO calculations[35] for $[Hg(SCH_3)_2]$ and $[Hg(SCH_3)_2(S\text{-}tert\text{-}C_4H_9)]^-$, as well as those obtained from the X-ray analyses of $[Hg(SCH_3)_2]$ and $[Hg(S\text{-}tert\text{-}C_4H_9)^3]^-$.[15f] The raw EXAFS functions and the FTs of the NaS-*tert*-C_4H_9/**2** = 1/1 and **2** in DMF were best explained by assuming Debye−Waller-like factors (σ) of 0.077 and 0.061 Å, respectively, for the first shell. These σ values are almost the same as those obtained from the analyses by curve fitting and ratio methods (Table 1). The results of FEFF calculations are summarized in Table 2. As shown in Figure 4a−d, the EXAFSs and FTs derived from these best fit calculations (dashed lines) based on the $Hg(SC)_3$ and $Hg(SC)_2$ core structures in Figure 5 reproduce the raw EXAFSs and FTs of the NaS-*tert*-C_4H_9/**2** = 1/1 and **2** solutions (solid lines) well. The change in coordination geometry from **2** to **1** (NaS-*tert*-C_4H_9/**2** = 1/1) is seen in the higher shells (nlegs = 3 in Table 2) of the radial distribution functions. In the region indicated by the double-headed arrow in Figure 4d, the FT of **2** exhibits a peak at 4.66 Å originating from the multiple scattering of the four-leg path, Hg → S → Hg → S → Hg, whereas the FT of the NaS-*tert*-C_4H_9/**2** = 1/1 solution shows almost no peak attributable to multiple scattering (see Figure 4b and Table 2).[36] The FEFF simulations of the solid samples of $[Hg(SCH_3)_2]$[15b,c] and $(NEt_4)[Hg(S\text{-}tert\text{-}C_4H_9)_3]$[15e,f] also reproduced their raw EXAFSs and FTs well and exhibited the same difference in higher shells as in the cases of NaS-*tert*-C_4H_9/**2** = 1/1 and **2** solutions.

Table 2. Results of FEFF Calculations[a] for [(Boc-CPLC-C-OMe)Hg][b] and [(Boc-CPLC-OMe)(S-*tert*-C_4H_9)Hg]$^-$ [b]

core	no.[c]	σ	%[d]	deg	nlegs	r_{eff}
$Hg(SC)_2$[e]	1	0.0606	100.00	2.00	2	2.3300
	2	0.1063	26.09	2.00	2	3.3548
	3	0.1006	23.41	4.00	3	3.7066
	4	0.1057	7.23	2.00	4	4.0584
	5	0.0855	6.64	2.00	3	4.6600
	6	0.0855	17.50	2.00	4	4.6600
$Hg(SC)_3$[f]	1	0.0766	100.00	3.00	2	2.4240
	2	0.1335	8.79	1.00	2	3.4816
	3	0.1335	8.76	1.00	2	3.4856
	4	0.1335	8.71	1.00	2	3.4915
	5	0.1266	7.77	2.00	3	3.8370
	6	0.1266	7.66	2.00	3	3.8438
	7	0.1266	7.70	2.00	3	3.8458

[a] FEFF 6.01.[28] [b] See the figure caption for Figure 5 about the geometries of the cores used in FEFF. [c] Path number. [d] Curved wave amplitude ratio. [e] SO^2 = 0.85; T = 295 K; Debye temperature Θ_D = 575 K; energy zero shift, ΔE_r = 7.0 eV, ΔE_i = −0.5 eV. [f] SO^2 = 0.85; T = 295 K; Debye temperature Θ_D = 450 K; ΔE_r = 7.0 eV, ΔE_i = −0.5 eV.

The Hg−S bond length obtained from the NaS-*tert*-C_4H_9/**2** = 1/1 solution (2.42 ± 0.03 Å) is not close to that of the starting material, **2** (2.33 Å),[19] but close to the 2.442 Å of the trigonal planar $(NEt_4)[Hg(S\text{-}tert\text{-}C_4H_9)_3]$.[15f] On the other hand, the observed Debye−Waller-like factors of this solution, σ = 0.0773 Å (from curve fitting and ratio methods; Table 2) and σ = 0.0766 Å (from the FEFF calculation), are larger than that obtained from the FEFF simulation of the experimental data of the solid state $(NEt_4)[Hg(S\text{-}tert\text{-}C_4H_9)_3]$, σ = 0.0601 Å. The reason for this discrepancy in σ can be explained by the static disorder[24] accompanying the ligand exchange reaction of $[S\text{-}tert\text{-}C_4H_9]^-$ in the NaS-*tert*-C_4H_9/**2** = 1/1 solution or by the fact that we are observing a mixture of **1** and **2** in the NaS-*tert*-C_4H_9/**2** = 1/1 solution, which contains more than 90% of **1** and less than 10% of **2**.

As a consequence, the results of the EXAFS analyses, referring to those of UV−vis and NMR experiments, strongly indicate that the Hg^{2+} of the dominant species generated in the NaS-*tert*-C_4H_9/**2** = 1/1 solution is trigonally coordinated but not linear.

NMR. We determined the structure of proline as *trans*[37] from the ^{13}C chemical shifts of Pro C_β (29.561 ppm) and Pro C_γ (24.405 ppm).

(34) On the other hand, EXAFS analyses for the NaS-*tert*-C_4H_9/**2** = 2/1 solution yielded ambiguous results. The coordination number that we obtained for this system was 4.0, which is inconsistent with the estimation from the XANES comparison mentioned above, whereas the Hg−S bond length was 2.441 Å, which is in agreement with those of the trigonal planar complexes such as $(NEt_4)[Hg(S\text{-}tert\text{-}C_4H_9)_3]$.[16e,f] As will be mentioned later, this is probably because the solution is an equilibrium mixture consisting of **1** and some other unknown Hg−S compounds.

(35) MOPAC Version 6: Stewart, J. J. P. *QCPE Bull.* **1989**, *9*, 10; revised as Version 6.01 by K., Nishida, of Fujitsu Co., Ltd., for UNIX Sun SPARK station and revised as Version 6.02 by the present authors for Silicon Graphics IRIS 4D.

(36) Generally speaking, multiple scattering intensity is strongest when ∠B−A−B = 180°, where A is the X-ray scatterer and B is the backscatterer, and decays as the bite angle decreases.

(37) Würthrich, K. *NMR in Biological Research: Peptides and Proteins*; North-Holland Publishing Co.: Amsterdam, 1976.

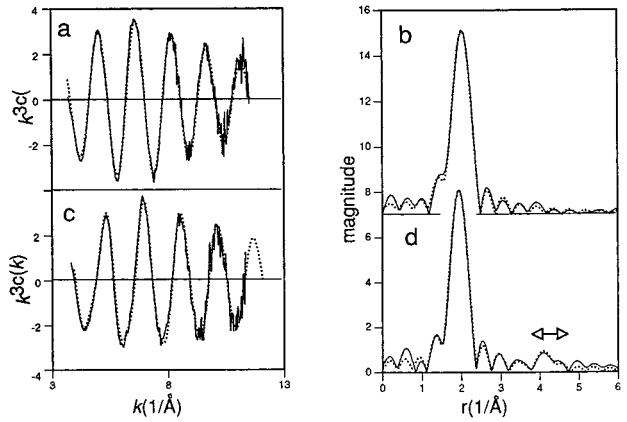

Figure 4. (a) Solid line, raw EXAFS function of NaS-*tert*-C_4H_9/**2** = 1/1; dashed line, EXAFS function for C_{3h} $Hg(SC)_3$ core (see Figure 5b) calculated using FEFF 6.01.[29] (b) Solid line, the Fourier transform of the Hg L(III) edge EXAFS of NaS-*tert*-C_4H_9/**2** = 1/1 in DMF (k = 3.8−11.6 Å^{-1}); dashed line, the Fourier transform of the dashed line in Figure 4a (k = 3.8−11.6 Å^{-1}). (c) Solid line, the raw EXAFS function of **2** in DMF; dashed line, the EXAFS function for C_{2h} $Hg(SC)_2$ core (see Figure 5a) calculated by FEFF 6.01.[29] (d) Solid line, the Fourier transform of the Hg L(III) edge EXAFS of **2** in DMF (k = 3.95−12.05 Å^{-1}); dashed line, the Fourier transform of the dashed line in Figure 4c (k = 3.95−12.05 Å^{-1}).

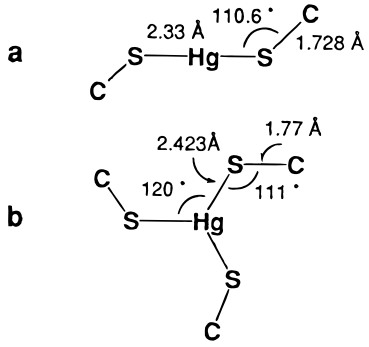

Figure 5. (a) $Hg(SC)_2$ and (b) $Hg(SC)_3$ cores used in FEFF calculations. The geometry of $Hg(SC)_2$ was constructed with reference to the results of X-ray analysis.[15,c] The geometry of $Hg(SC)_3$ was constructed with reference to the results of EXAFS calculation (Table 1), the results of X-ray crystallography for $[Hg(S\text{-}tert\text{-}C_4H_9)_3]^-$ by O'Halloran's group,[15f] and the calculation using MOPAC 6.01.[35]

The linearity observed for the ^{1}H NMR signals of the NaS-*tert*-C_4H_9/**2** = 0−1.0 solutions (Figure 3) allows us to assign the prochiralities of the bands of Cys^1 $H_\beta{}^1$, Cys^1 $H_\beta{}^2$, Cys^4 $H_\beta{}^1$, Cys^4 $H_\beta{}^2$, Pro $H_\beta{}^1$, Pro $H_\beta{}^2$, Pro $H_\delta{}^1$, and Pro $H_\delta{}^2$ in the NaS-*tert*-C_4H_9/**2** = 1/1 sample to be Cys^1 $H_\beta{}^{1S}$, Cys^1 $H_\beta{}^{2R}$, Cys^4 $H_\beta{}^{1S}$, Cys^4 $H_\beta{}^{2R}$, Pro $H_\beta{}^{1S}$, Pro $H_\beta{}^{2R}$, Pro $H_\delta{}^{1R}$, and Pro $H_\delta{}^{2S}$, respectively, on the basis of the band assignment for **2**.[19] The prochiralities of the Pro H_β were also determined from the NOESY spectrum by comparing the off-diagonal peak intensities between Pro $H_\beta{}^{1,2}$ and Pro H_α. The results agreed with those determined from the linearity in Figure 3.

Temperature-dependent chemical shifts (δ/T; ppm·deg^{-1}) were observed to reveal shielding of the amide protons (H_Ns) from the solvent.[11,37] δ/T values for the H_Ns of the 1/1 solution were found to be linear over the temperature range of −50 to 25 °C, and were 0.0103 for Cys^1 H_N, 0.0010 for Leu H_N, and 0.0011 for Cys^4 H_N. Referring the results of the δ/T-NOE experiments on the Cys-Pro-X-Cys disulfides carried out by Balaram's group,[11] the results seem to suggest strong shielding of the Leu H_N and Cys^4 H_N from the solvent.[38] This leads to the possibility that these protons are hydrogen bonded to some atoms. The heteronuclear multiple-bond correlation (HMBC)[23] experiment were examined to find the CO groups as the hydrogen bonding partners of Leu H_N and Cys^4 H_N but failed. This result suggests that the sulfur atoms of the Cys^{1,4} are the hydrogen-bonding partners of the Leu H_N and Cys^4 H_N.

Twenty one ^{1}H−^{1}H distances for the main chain loop, including the Cys^1 S−Hg−Cys^4 S bridge, were estimated from the off-diagonal peak volumes of the PS NOESY and ROESY spectra using the ^{1}H−^{1}H distance (1.752 Å) of the geminal Pro $H_{\beta(\text{or }\delta)}$ as the internal standard.[39] The distances for vicinal H_N−H_α pairs were also estimated from their coupling constants, $^3J_{HN\alpha}$ = 9.5 Hz (Leu) and 9.0 Hz (Cys^4), using the equation, $^3J_{HN\alpha}$ = 6.4 $\cos^2\theta$ − 1.4 $\cos\theta$ + 1.9, formulated by Pardi et al. for bovine pancreatic trypsin inhibitor, where θ is the dihedral angle $\angle H_N\text{−}N\text{−}C\text{−}H_\alpha$.[40] The H_α−H_β distances for Cys^{1,4} derived from $^3J_{H_\alpha H_\beta}$ using the Kopple equation[41] were not adopted here due to the changes in the Cys^{1,4} S−Hg bonding nature. From these interproton distances, the lower and upper limits, J_{ij} and U_{ij}, respectively, of the distances for the proton pairs were estimated and listed in Table 3.[42]

Table 3. Atom Distance Constraints and the Violations for $[(\text{Boc-Cys-Pro-Leu-Cys-OMe})(S\text{-}tert\text{-}C_4H_9)Hg]^-$

origin[a]	atom 1[b]	atom 2[b]	L[c]	U[d]	V[e]
A	Hg	Cys^4 S	2.420	2.426	−0.024
A	Cys^1 S	Cys^4 S	4.282	4.288	−0.144
A	Cys^1 S	S-*tert*-C_4H_9	4.298	4.304	−0.267
B	Leu H_N	Pro $H_\delta{}^{2,\text{pro-S}}$	1.99	3.12	0.000
B	Leu H_N	Cys^4 H_α	2.87	4.79	0.095
B	Leu H_N	Leu H_α	2.78	2.99	0.000
B	Leu H_N	Cys^4 H_N	1.32	1.85	0.342
B	Cys^4 H_N	Cys^4 $H_\beta{}^{1,\text{pro-S}}$	1.62	2.46	0.347
B	Cys^4 H_N	Cys^4 $H_\beta{}^{2,\text{pro-R}}$	2.34	3.79	0.636
B	Cys^4 H_N	Cys^4 H_α	2.70	2.99	0.000
B	Cys^4 H_N	Leu H_α	2.70	3.90	0.000
B	Cys^1 H_α	Cys^1 $H_\beta{}^{1,\text{pro-S}}$	2.23	2.60	−0.017
B	Cys^1 H_α	Cys^1 $H_\beta{}^{2,\text{pro-R}}$	2.50	2.93	0.000
B	Cys^1 H_α	Pro $H_\delta{}^{1,\text{pro-R}}$	1.69	2.60	0.000
B	Cys^1 H_α	Pro $H_\delta{}^{2,\text{pro-S}}$	1.76	2.73	0.000
B	Cys^4 H_α	Cys^4 $H_\beta{}^{1,\text{pro-S}}$	2.40	2.80	0.000
B	Cys^4 H_α	Cys^4 $H_\beta{}^{2,\text{pro-R}}$	2.23	2.60	−0.056
B	Pro H_α	Pro $H_B{}^{1,\text{pro-S}}$	2.23	2.60	−0.057
B	Pro H_α	Pro $H_\beta{}^{2,\text{pro-R}}$	2.23	2.60	0.010
B	Pro $H_\delta{}^{2,\text{pro-S}}$	Pro $H_\beta{}^{1,\text{pro-S}}$	2.32	3.75	0.703
B	Pro $H_\delta{}^{2,\text{pro-S}}$	Cys^1 $H_\beta{}^{1,\text{pro-S}}$	2.55	4.25	0.000
B	Cys^1 H_N	Cys^1 $H_\beta{}^{1,\text{pro-S}}$	2.87	4.79	0.000
B	Cys^1 H_N	Cys^1 $H_\beta{}^{2,\text{pro-R}}$	1.95	3.10	0.000
B	Cys^1 H_N	Cys^1 H_α	2.50	2.93	0.000

[a] A by means of EXAFS, B by NOESY. [b] The superscripts on the right shoulders of $H_{\alpha-\delta}$ represent the ^{1}H NMR numbering labeled from high field to low field. [c] Lower limit (Å). [d] Upper limit (Å). [e] + and − denote the violations from the upper and lower limits, respectively.

(38) The deuterium exchange experiment using CH_3OD in DMF-d_7 afforded no obvious result as for **2**.

(39) The ROESY spectrum was obtained to avoid erroneous information due to the rapid exchange between **1** and **2**. Both spectra were basically in good agreement with each other.

(40) Pardi, A.; Billeter, M.; Würthrich, K. *J. Mol. Biol.* **1984**, *180*, 741.
(41) Kopple, K. D.; Wiley, G. R.; Tauke, R. *Biopolymers* **1973**, *12*, 627.

(42) For the vicinal H_N−H_α distances obtained from $^3J_{HN\alpha}$, we allotted ±0.2 Å latitudes within the theoretical distance (2.19−3.00 Å). The H_α−H_β distances of Cys^{1,4} were classified as 2.23−2.60, 2.40−2.80, and 2.60−3.04 Å according to their ROESY cross peak volumes, i.e., raw distances. The use of coupling constants was impossible because the methylene groups were under the influence of Hg−S bonds. We determined the lower and upper limits of the other interproton distances, using the equations $L = 1.0r - 0.6$ Å and $U = 1.9r - 1.8$ Å, respectively, where r is the raw distance. The other two types of latitudes were examined: (1) the lattitude described in the previous paper[19] and (2) the latitude described by Peishoff et al. (Peishoff, C. E.; Bean, J. W.; Kopple, K. D. *J. Am. Chem. Soc.* **1991**, *113*, 4416). The former was coarse, whereas the latter was so narrow that it gave a poor convergence yield.

We observed 25 of $^1H-^1H$ pairs for the peptide side chains in the ROESY spectrum. They were not included in Table 3 except for those including Pro $H_{\beta,\delta}$ because of different correlation time, τ_c. However, we emphasize here that we observed in the ROESY spectrum weak off-diagonal peaks with respect to the Cys[4]−OMe: Cys[4] $H_\beta{}^{1S}$−Cys[4] OMe and Cys[4] $H_\beta{}^{2R}$−Cys[4] OMe, Leu H_α−Cys[4] OMe, Boc−Cys[4] OMe, and S-*tert*-C_4H_9−Cys[4] OMe.

Structural Analysis

Data Processing and Calculations. Structural analyses were performed using three different calculations: DG, MD, and RMD. The DG calculation was first performed to obtain experimentally the most likely structure from the distance information (Table 3) obtained from EXAFS and ROESY experiments. The best structure obtainable from DG calculations is not always best in energy. Thus, as the second step, we performed the MD calculations. MD calculation, however, includes a bias error due to the incompleteness in force field and treatment of the solvent. Thus, as the third step, RMD based on the experimental constraint was carried out, and the results from the DG, MD, and RMD calculations were compared.

The program CADMOS[19,43] was used for the DG calculation. CADMOS is based on the algorithm of the variable target function (TF) method.[44] It combines the atom distance information from XAFS experiments (for inorganic center) and NOESY or ROESY experiments (for peptide) into the TF as weighted terms. The program is equipped with the optimization module based on an automated simulated annealing[45] with discrete temperature decrement.[46] CADMOS calculates the coordination center by two ways: (1) as a new amino acid residue; (2) as an independent molecule with another Cartesian coordinates. The value of TF, ϵ, to be optimized is defined by eq 2 which includes

$$\epsilon = F_1 + \omega F_2 + \gamma F_{\mathrm{vdw}} \qquad (2)$$

three kinds of terms, F_1, F_2, and F_{vdw}. ω and γ are the parameters for the relative weights of F_2 and F_{vdw}. F_1 includes subterms belonging to the rough distance category, in which NOESY distance information is included, whereas F_2 corresponds to the more accurate category dealing with XAFS data. The other distances between the coordination center and the peptide unit which can be obtained from XAFS data and the structure of the coordination unit are also classified into category F_2. F_{vdw} deals with the Lennard−Jones repulsive cores for pairs of atoms at contact radii.

For every DG calculation, we prepared 150 random conformers and minimized their TF. Then, we selected the 50 structures with highest ϵ and employed them in the analyses.

MD calculations were performed by AMBER version 3.0 revision A using AMBER 4.1 force fields for the peptide moiety.[47] The force fields for the mercury centers were prepared by the procedure described below. RMD calculations were carried out by replacing the harmonic term[48] in AMBER with the quatic term[49] to facilitate the convergence of the internuclear distances into the constraint at the initial stage of the RMD calculation.

Concerning the (R)MD parameters of the $HgS_{2,3}$ centers, partial charges were obtained from the model calculation for $[Hg(SCH_3)_2(S\text{-}tert\text{-}C_4H_9)]^-$ using MULLIK of MOPAC 6.01.[35,50] These charges were then modified for **1** so that the total charge equaled −1 and the charges on the main chain (AMBER 4.1 force field) were conserved (Supporting Information). The force constants of the HgS_3 core of **1** were derived from the model calculation for $[Hg(SCH_3)_3]^-$ using AMBER and MOPAC, in such a way that the AMBER simulation reproduced the frequencies from MNDO structural optimization (Supporting Information). This is based on the fact that the frequencies obtained by MNDO calculation, experimental measurements, and AMBER simulation for $[Hg(SCH_3)_2]$ were in good agreement (Supporting Information).[19] The van der Waals radii and the depths of the van der Waals wells were set to be the same as those in our previous paper.[19]

Each MD calculation was carried out under vacuum with a step size of 1 fs by employing the program SHAKE.[51] Stepwise annealing was applied with decreasing temperature from 700 to 600, 500, and 400 K (with 5000 steps for each), and the MD was finally equilibrated at 300 K using 40 000 steps. Each temperature was held by the hypothetical thermal exchange between the metallopeptide and the heat bath.[52] The time constant for heat bath coupling was set at 0.1 ps. For each MD calculation, 5000 conformers were randomly generated as the starting points and minimized by molecular mechanics calculations. Among these, 100 of the lowest-energy structures were subjected to the MD calculations without several molecules. The distance-dependent dielectric constant $\epsilon_{ij} = 1.5r_{ij}$ was used according to the previous paper.[19] The averaged structures and the energies for the final 10 ps were used for discussion. RMD calculations were performed by basically the same procedure. The scaling factor for 1−4 interactions and 1−4 Coulombic interactions was set at 2.0.

Molecular graphics images were produced using the MidasPlus software system version 1.5.[53] All calculations, DG, MD, RMD, and EXAFS analyses, were performed on a HP 9000/735 workstation.

Results of the Distance Geometry Calculation. In the DG calculation, we first surveyed the hydrogen-bonding partner of the Leu H_N and Cys[4] H_N, adding hypothetical hydrogen bonds to the constraints, and compared the histograms of the 50 structures with highest ϵ. Two hydrogen-bonding schemes were found to give the lowest distributions of the target function: type DG-a; Cys[1] S−Leu H_N, Cys[1] S−Cys[4] H_N; type DG-b, Cys[1] S−Leu H_N, Cys[1] CO−Cys[4] H_N. Figure 6-1 shows the 40 selected structures belonging to the DG-a and DG-b types. The structures in this figure are superimposed at the 7 atoms of the

(43) F_1, F_2, and F_{vdw} are given by F_1, $F_2 = \sum'(U_{ij}^2 - r_{ij}^2)^2/N_{1,2} + \sum'(r_{ij}^2 - L_{ij}^2)^2/N_{1,2}$ and $F_{\mathrm{vdw}} = \sum[S_i + S_j)^2 - r_{ij}^2]^2$, respectively, where r_{ij} is the distance between the ith and jth atoms belonging to a structure generated in computation. U_{ij} and L_{ij} are the constraints for r_{ij} ($L_{ij} < r_{ij} < U_{ij}$). S_i and S_j are the repulsive core radii of atoms i and j. $\sum'$ means that the summation is feasible only for the atom pairs which do not satisfy the constraint $L_{ij} < r_{ij} < U_{ij}$ or $r_{ij} > (S_i + S_j)$.

(44) Braun, W.; Go, N. *J. Mol. Biol.* **1985**, *186*, 611.
(45) Metropolis, N.; Rosenbluth, A.; Teller, A.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087. Kirkpatrick, S.; Gelatt, C. D., Jr.; Vecchi, M. P. *Science* **1983**, *220*, 671.
(46) Otten, R. H. J. M.; Ginneken, L. P. P. P. *Research*; IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 1984; RC-10861.
(47) AMBER 3.0, Revision A: Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
(48) Pearlman, D. A.; Kollman, P. A. *J. Mol. Biol.* **1991**, *220*, 457.
(49) Nilges, M.; Gronenborn, A. M.; Brünger, A. T.; Clore, G. M. *Protein Eng.* **1988**, *2*, 27.
(50) The result of the PM3 calculation was not adopted, because the structural optimization for $[Hg(SCH_3)_2]$ using this Hamiltonian afforded ∠Hg−S−C = 90°. This is different from the result of 111° in the X-ray analysis of this compound.[20b,c] In order to obtain satisfactory results for the parameters of HgS compounds, it is necessary to achieve relativistic ab initio calculations. We did not carried out this calculation because of the long cpu time. Furthermore, the basis sets for Hg are at present not believable.
(51) Gunsteren, W. A.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.
(52) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(53) Ferrin, T. E.; Fuang, C. C.; Jarvis, L. E.; Langridge, R. The MIDAS display system. *J. Kol. Graphics* **1988**, *6*, 13.

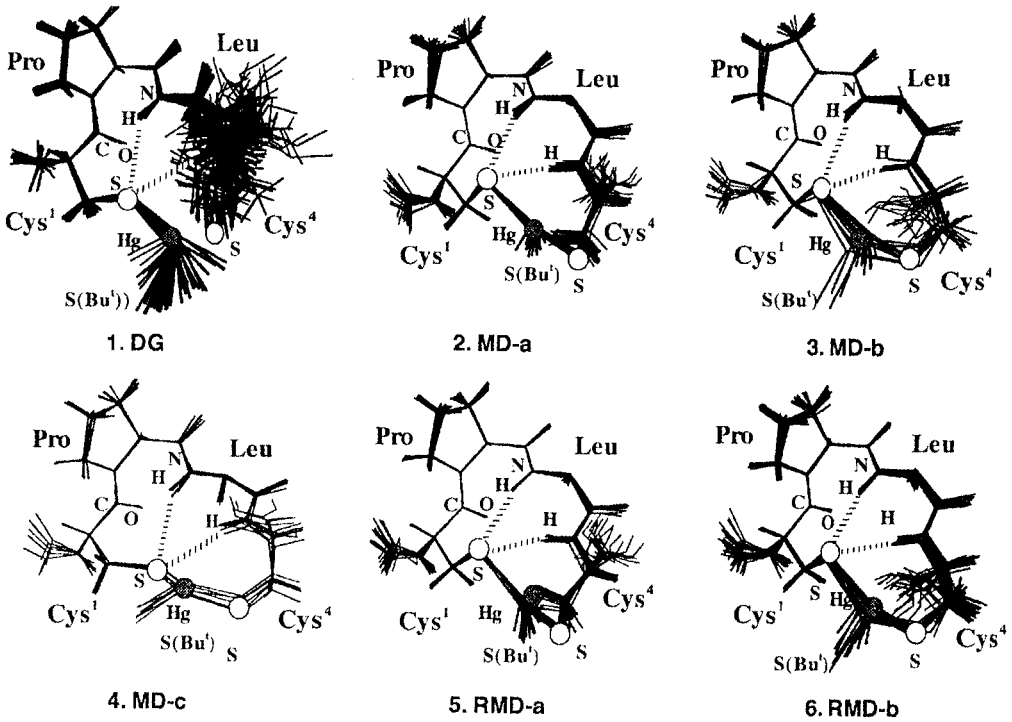

Figure 6. Converged structures of [(Boc-Cys-Pro-Leu-Cys-OMe)(S-*tert*-C_4H_9)Hg]$^-$, **1**, superimposed at Cys^1 C_β, Cys^1 C_α, Cys^1 C, Pro N, Pro C_α, Pro C, and Leu N. The dashed lines represent the hydrogen bonds. 1. DG: 35 models from the DG calculation. For clarity, the Cys^4 S−Hg bonds are not shown in the figure (rmsd ≤ 0.0898 Å for the seven atoms). 2. MD-a: 21 models from the MD calculation (rmsd ≤ 0.0771 Å). 3. MD-b: 13 models from the MD calculation (rmsd ≤ 0.1086 Å). 4. MD-c: 6 models from the MD calculation (rmsd ≤ 0.10024 Å). 5. RMD-a: 18 models from the RMD calculation (rmsd ≤ 0.06239 Å). 6. RMD-b: 22 models from the RMD calculation (rmsd ≤ 0.08881 Å).

N-terminal main chains (Cys^1 S, Cys^1 C_β, Cys^1 C_α, Cys^1 C, Pro N, Pro C_α, Pro C, and Leu N) with rmsd ≤ 0.3410 Å. As shown here, the structural displacements at the C-termini were too large to distinguish the two structural types at the C-termini.

Results of Molecular Dynamics. The MD calculation afforded three types of stable structure: MD-a (Figure 6-2), MD-b (Figure 6-3), and MD-c (Figure 6-4). These three gave almost the same histogram in potential energy distribution and exhibited two common hydrogen bonds: Cys^1 S−Leu H_N and Cys^1 S−Cys^4 H_N. Among these structures, MD-b is the isostructure of the ZF[3] and Rub[4] cores. The structure of MD-a is different from that of MD-b in the rotation of Cys^4 C_α−Cys^4 C_β, as well as in the direction of the C-terminus, Cys^4 C_α−Cys^4 C(O). The structure of MD-c differs from those of MD-a, MD-b, DG-a, and DG-b but shows the same Cys^1 C_α−Cys^1 C_β rotation as **2**[19] (Figure 7a).

A long-term MD calculation that extended to 10 ns (300 K) was then carried out, starting from the average of the 21 structures in MD-a. Figure 8 shows the time courses of the potential energy of this calculation (Figure 8A), the Cys^1 χ_1 rotation (Figure 8B), the Cys^4 χ_1 rotation (Figure 8C), and the Cys^1 S−Cys^4 H_N hydrogen-bonding distance (Figure 8D). Here and hereafter, χ_1 and χ_2 are defined as $Cys^{1,4}$ N−$Cys^{1,4}$ C_α−$Cys^{1,4}$ C_β−$Cys^{1,4}$ S and $Cys^{1,4}$ C_α−$Cys^{1,4}$ C_β−$Cys^{1,4}$ S−Hg^{2+}, respectively. As seen in the figure, the Cys^1 S−Cys^4 H_N hydrogen bond is conserved throughout the calculation, in spite of that the time courses of A, B, and C are discontinuous. The average structures extracted from the areas interposed by the discontinuity correspond to the three structural types, MD-a, MD-b, and MD-c. Throughout the conformational exchange, the rotations of Cys^4, Cys^4 χ_1, and Cys^4 χ_2 have almost no effect on the energy, whereas those of Cys^1, Cys^1 χ_1, and Cys^1 χ_2 cause a 10 kcal/mol difference in energy. As a consequence, we see that the Cys^4 part of **2** is flexible. On the other hand, we find that MD-c is the computational grand minimum from this time course. Considering that the DG calculation afforded almost no conformations that exhibit the same Cys^1 C_α−Cys^1 C_β rotation in the lower TF region, MD-c structures should be excluded from the candidates for the actual structure. The MD-c structure probably originates from the bias error of the force field, especially that of the HgS_3 center, or the lack of solvent molecules in the MD calculation.

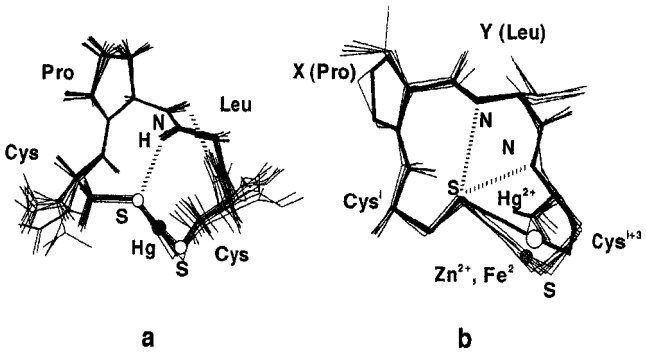

Figure 7. Superposition of the converged structures of **2**.[19] (b) Superposition of the best of the RMD-b type structures (the thick line) and the main chain loops of the 8 cores of Cys-X-Y-Cys/M^{2+} in proteins (thin lines), taken from the X-ray analyses for AT (M = Zn),[2] ZF (M = Zn),[3] and Rub (M = Fe).[4] The rmsd is less than 0.16316 Å for the 10 atoms of the main chain except for Cys^1 C_β, Cys^1 S, Hg, Cys^4 S, and Cys^4 C_β.

Results of Restrained Molecular Dynamics. The RMD calculations afforded three types of structure corresponding to those of the MD calculation: RMD-a, RMD-b, and RMD-c. Among these, RMD-b was the major type (Cartesian coordinates: Supporting Information). The RMD-c structures (11 models, rmsd ≤ 0.079 Å) were excluded for the same reason as mentioned in the section on the MD calculation.[54] The other two types are shown in Figure 6 as those superimposed on the N-terminal 7 atoms: RMD-a, Figure 6-5; RMD-b, Figure 6-6.

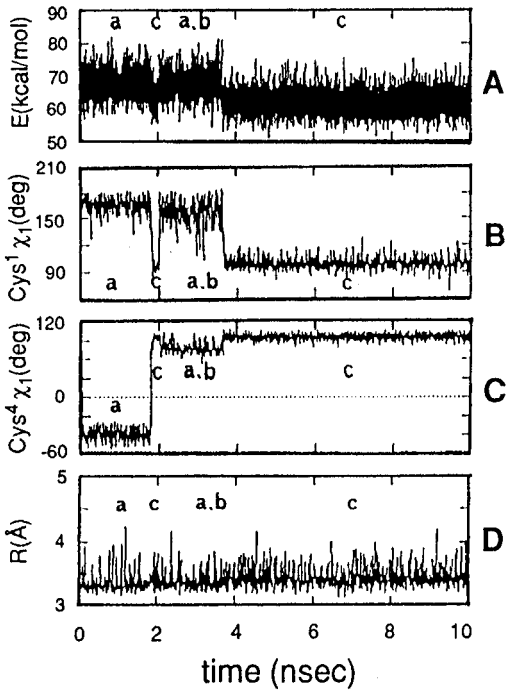

Figure 8. Conformational exchange among MD-a, MD-b, and MD-c, shown by the time courses of (A) the potential energy, (B) Cys^1 χ_1 defined as $\angle N-C_\alpha-C_\beta-S$, (C) Cys^4 χ_2 defined as $\angle C_\alpha-C_\beta-S-Hg$, and (D) the distance R (Å) between Cys^1 S and Cys^4 N. The structural types are indicated by a, b, and c instead of PMD-a, PMD-b, and PMD-c, respectively.

Discussion

Actual Structure. The ^1H NMR spectra were completely homogeneous over the temperature range from −55 to 25 °C. On the other hand, we observed weak NOEs between Leu $H_\alpha-Cys^4$ OMe, as well as Boc−Cys^4 OMe and S-*tert*-$C_4H_9-Cys^4$ OMe. These results lead to two possibilities. (1) There is a dominant species in the NaS-*tert*-C_4H_9/**2** = 1/1 solution, and the structure is the intermediate of those of (R)MD-a and (R)MD-b. (2) There is a rapid equilibrium among plural species, one of which is structurally similar to (R)MD-b. The latter image agrees with the results of the (R)MD calculations, which indicated that the conformational change between (R)MD-a and (R)MD-b causes almost no change in energy despite the differences in the Cys^4 χ_1 and Cys^4 χ_2 rotations, if the energy barrier between (R)MD-a and (R)MD-b is low. The former leads to the inaccuracy of the (R)MD calculations, and the energy discrepancy between RMD-a and RMD-b is truly large. It is impossible to judge whether we observed (R)MD-b or a mixture of (R)MD-a and (R)MD-b in the NMR spectra of NaS-*tert*-C_4H_9/**2** = 1/1 in DMF-d_7 from the (R)MD calculations done under vacuum conditions.

Finally, a MD calculation including 128 molecules of DMF was carried out starting from the best (in constraint and energy) structure of RMD-b (Figure 7b; thick line). This calculation revealed no structural changes up to 1 ns as observed in the calculation done under vacuum conditions. Consequently, we can say that the actual structure of the dominant species in the NaS-*tert*-C_4H_9/**2** = 1/1 solution is close to (R)MD-b.

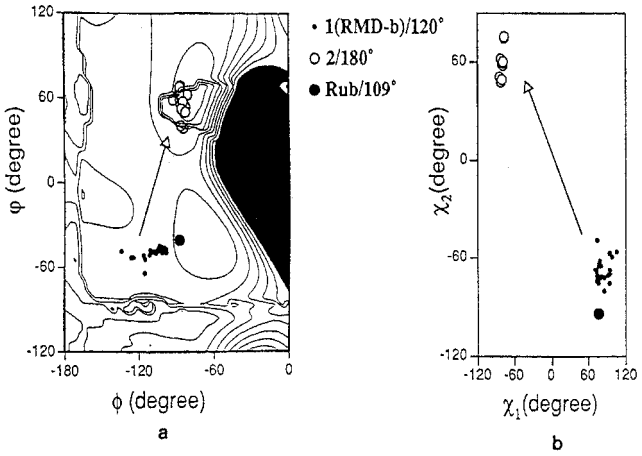

Figure 9. (a) Ramachandran plots of the Leu (ϕ, φ) for **1** and **2** and Val (ϕ, φ) of the Rub core site on the potential surface of AcNH-Leu-CONHMe. (b) Cys^4 (χ_1, χ_2) plots for **1** and **2** and the Rub core site. The shifts of the distribution areas for (ϕ, φ) and (χ_1, χ_2) are shown by arrows. The potential energy surface of the Leu (ϕ, φ) was calculated for every 5° grid of (ϕ, φ) on an AMBER 4.1A force field by minimizing the side chain energy. The contour lines are shown for every 2 kcal mol^{-1} from the lowest energy. The high-energy area is shaded for simplicity.

Structural Comparison. The (R)MD-b structure is common to Cys^i-X-Y-Cys^{i+3}/M^{2+} cores in proteins.[2−4] The best RMD-b structure is in close agreement with that of the Cys^i-X-Y-Cys^{i+3}/M^{2+} in proteins (Figure 7b) despite the differences in metal ion species (rmsd ≤ 0.163 Å for the 10 atoms of the main chain loop). On the other hand, the structure differs from that of **2** (Figure 7a) in terms of the hydrogen-bonding scheme, the dihedral angles at Pro (ϕ, φ), Leu (ϕ, φ), $Cys^{1,4}$ (χ_1, χ_2), and the direction of the C-terminus. This originates from the difference in bite angle. The bite angle discrepancy between compound **1** ($\angle S-Hg-S = 120°$) and the cores in biological systems ($\angle S-Zn(Fe)-S \cong 110°$) is small, whereas that between **1** and **2** is large.

Figure 9a shows the distribution of Leu (ϕ, φ) of **1** (RMD-b structures) and **2**[19] as well as Val (ϕ, φ) of the Rub core site plotted on the contour map of the Leu (ϕ, φ) potential surface calculated using AMBER 4.1A. The plots for Pro (ϕ, φ) are not shown here because they were found at the same locations as those in proteins (ϕ, −90 to −50°; φ, −50 to 0°). As seen in the figure, the Leu (ϕ, φ) sets of **1** are distributed in (ϕ, −126.1 to −98.1°; φ, −54.3 to −44.8°) for RMD-b. This deviates negatively from the ordinary α helix region (ϕ, −80 to −60°; φ, −50 to −30°) in ϕ[55] and is close to the Val (ϕ, −87.4°; φ, −40.8°) of the Rub core site, as expected from the superposition in Figure 7b. The Leu (ϕ, φ) does not overlap, however, with the (ϕ, −92.5 to −81.0°; φ, 38.3 to 68.2°) of the Leu (ϕ, φ) area of **2**, which lies in the general turn area.

The effect of trigonal planar Hg^{2+} on the structure of the proximal part is represented by the (χ_1, χ_2) map of the $Cys^{1,4}$ $C_\alpha-C_\beta$ dihedral angles (Figure 9b). The (χ_1, χ_2) of Cys^1 of RMD-b is distributed in (χ_1; 158.5 to 178.6°, χ_2; −167.1 to −155.6°), and that of Cys^4 lies in (χ_1, 74.8 to 106.5°; χ_2, −80.8 to −50.0°), whereas those of **2** are distributed in (χ_1, −30.9 to 3.0°; χ_2, −48.6 to −65.1°) for Cys^1 and (χ_1, −83.0 to −76.4°; χ_2, 47.2 to 75.4°) for Cys^4. Despite the difference in the

(54) In the search using the DG calculation, we assumed possible hydrogen-bonding schemes on the basis of on NMR experiments. Although this search failed to converge the structures to a well-defined one because of the insufficiency of the distance information for the C-terminal half, the obtained N-terminal halves were common in the main chain folding and the hydrogen-bonding Cys^1-*S*-Leu-H_N. The same type of N-terminal half is also observed for the structures from the MD, MD-a (Figure 6-2), and MD-b (Figure 6-3). This agreement between the most probable structures constructed on the basis of experimental constraint and a priori force field markedly increases the validity of the N-terminal halves of the RMD structures, RMD-a (Figure 6-5) and RMD-b (Figure 6-6).

(55) Richardson, J. S.; Richardson, D. C. In *Prediction of Protein Structure and the Principles of Protein Conformation*; Fasman, G. D., Ed.; Plenum Press: New York, 1989; p 1.

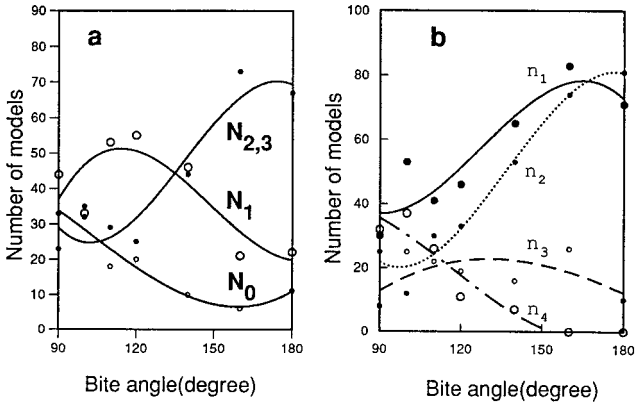

Figure 10. Rise and decay of the structures depending on the hypothetical ∠S−Hg−S bite angles of 100, 110, 120, 140, 160, and 180°. Key: (a) Numbers of structures involving multiple hydrogen bonds ($N_{2,3}$), one hydrogen bond (N_1), and no hydrogen bond (N_0). (b) Numbers of structures involving Cys[1] S−Leu H_N (n_1), Pro CO−Cys[4] H_N (n_2), Cys[1] CO−Cys[4] H_N (n_3), and Cys[1] S−Cys[4] H_N (n_4).

rotations of these dihedral angles, the conformational energies of **1** and **2** are almost the same.

Effect of ∠S−Hg−S Bite Angle. The conformational difference between **1** and **2** suggests the importance of the ∠S−M−S bite angle in determining the peptide conformations of Cys^i-X-Y-Cys^{i+3}/M^{2+}. In order to clarify this point, we carried out a series of MD calculations in vacuo, varying the ∠S−Hg−S bite angles as 90, 100, 110, 120, 130, 140, 160, and 180°. In the course of this calculation, the atomic charges and force constants were maintained at the same values as those for **2**. For each bite angle, 100 MD structures were obtained in the same manner as that described in the section for MD and classified with respect to the hydrogen-bonding scheme.

The comparison of potential energy histograms for the 8 bite angles (Supporting Information), using 100 models for each, showed that the bite angle change had very little effect on the value of the lowest energy of the distribution.[56] However, the histograms showed two common peaks: A and B. Peak A increased with increasing bite angle, whereas peak B, which was about 10 kcal/mol higher than A, decreased with the angle. In order to obtain insight into the origin of these two peaks, we surveyed the molecular structures with respect to hydrogen bonds. Figure 10a shows the changes in the number of structures including more than two hydrogen bonds ($N_{2,3}$), one hydrogen bond (N_1), and no hydrogen bond (N_0). The energy histogram for each $N_{2,3}$, N_1, and N_0 clarified that the common peak B corresponds to the structures without hydrogen bonds (N_0).

Figure 10b shows the changes in the number of structures with respect to the hydrogen bonds; Cys[1] S−Leu H_N (n_1), Pro CO−Cys[4] H_N (n_2), Cys[1] CO−Cys[4] H_N (n_3), and Cys[1] S−Cys[4] H_N (n_4). From inspection of this figure, as well as Figure 10a, we conclude several points. (1) The critical point lies between 130 and 150°, around which N_1 exchanges with $N_{2,3}$ (Figure 10a). On the other hand, the number of structures with no hydrogen bonds, N_0 in Figure 10a, decreases monotonically. As a result, the structural diversity of the peptide is largest around 110°, which corresponds to the tetrahedral geometry of the central metal ion. (2) The Cys[1] S−Leu H_N and Pro CO−Cys[4] H_N hyrogen bonds appear throughout the bite angle change, and their numbers increase with increasing bite angle. The Cys[1] S−Cys[4] H_N hydrogen bond, however, decreases monotonically. The appearance of Cys[1] CO−Cys[4] H_N is almost constant. Cys[1] S−Leu H_N and the conformation of the N-terminal half are conserved for the structures with more than two hydrogen bonds throughout the bite angle change. As a result, with increasing bite angle the hydrogen-bonding partner of Cys[4] H_N is altered from Cys[1] S to Pro CO via Cys[1] CO for the structures with more than two hydrogen bonds. This accompanies the rotations of Leu (ϕ, φ), Cys[4] (ϕ, φ), and Cys[4] (χ_1, χ_2), as well as the change of the C-terminal vector.

Fixation of Cys[1]-X-Y-Cys[4] in Actual Systems. The largest structural diversity of peptides at a bite angle of about 110° differs greatly from our original expectation that 110° is the angle that most efficiently stabilizes the Rub type structure for Cys[1]-Pro-Leu-Cys[4]/M^{2+} and that the conformation is the most robust at this angle. Our computational experiment on Cys[1]-Pro-Leu-Cys[4]/M^{2+}, although it ignores the third and fourth ligands as well as solvent effects, indicates that, statistically, a tetrahedral metal ion stabilizes the conformation via the entropy effect and, dynamically, a tetrahedral metal ion allows the peptide to deal with external stress through the none-hydrogen-bonding conformations without consuming more than 10 kcal/mol of conformational energy. In other words, Cys^i-X-Y-Cys^{i+3} with tetrahedral M^{2+} has a high degree of freedom when it is free from the third and fourth ligands.

However, the actual Cys^i-X-Y-Cys^{i+3} is fixed. The conformational fixation of the actual Cys^i-X-Y-Cys^{i+3} is effected due to the existence of the third and fourth ligands which change the van der Waals interactions, electrostatic interactions, and local dielectric constant. In the case of **1**, the negative charge of the inorganic center strengthens the negativity of Cys[1,4] S. Therefore, the Cys[1] S−Cys[4] H_N hydrogen bond of **1** is stronger than that of the 110° in the above calculations. Among the cases of Cys^i-X-Y-Cys^{i+3}/M^{2+} (M = Zn, Fe) in biological systems, the inroganic centers, $[Zn(Cys)_4]^{2-}$,[3a,57] $[Zn(Cys)_3(His)]^-$,[58] $[Zn_2(Cys)_6]^{2-}$,[3f] and $[Fe(Cys)_4]^{2-}$,[4] are negatively charged similarly to $[Hg(S\text{-}Cys)_2(S\text{-}tert\text{-}C_4H_9)]^-$ of **1**. On the other hand, $[Zn(Cys)_2(His)_2]$[3c,3d,59] attains the same conformation in spite of its zero charge. This can be explained as follows. The lone pairs of the His imidazole groups increase the negative charge of the Cys^i S. In addition, the molecular model indicates the existence of steric hindrance between the imidazole and the side chain of the Y amino acid. This interaction pulls out M^{2+} from Cys^i-X-Y-Cys^{i+3} and rotates Leu (ϕ, φ). This allows hydrogen-bond formation between Cys^i S and Cys^{i+3} H_N.

Effect of C-Terminal Turning. The change of ∠S−Hg−S bite angle from 120 to 180° induces the rotation of the C-terminus of Cys^i-X-Y-Cys^{i+3}. In the Figures 1a and 7a, the C-terminus of **2** is directed to the right; on the other hand, the C-termini of the RMD-b structures and Cys^i-X-Y-Cys^{i+3}/M^{2+} (M = Zn, Fe) in proteins rotate about 90° from those of **2** as seen in the Figures 1b and 7b. The weak ROESY peaks observed for Leu H_α−Cys[4] OMe, Boc−Cys[4] OMe, and S-*tert*-

(56) We suggested in a previous paper that the conformation energy of the peptide in **2** (corresponds to 180°) is about 10 kcal/mol higher than that of the Rub core site (corresponds to 110°) in energy. However, in that estimation, the hydrogen bond Pro-*CO*-Cys[4]-*H_N* was ignored. Detailed comparison of the Leu (ϕ, φ), Pro (ϕ, φ), and Cys[1,4] (χ_1, χ_2) distributions among **1**, **2**, the Rub core site, and the cases in the calculation for the bite angle change indicated their equivalence in terms of the peptide conformational energies.

(57) Omichinski, J. G.; Clore, G. M.; Schaad, O.; Felsenfeld, G.; Trainor, C.; Appella, E.; Stahl, S. J.; Gronenborn, A. G. *Science* **1993**, *261*, 438.
(58) Summers, M. F.; South, T. L.; Kim, B.; Hare, D. R. *Biochemistry* **1990**, *29*, 329. Summers, M. F. *J. Cell. Biochem.* **1991**, *45*, 41.
(59) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635.

$C_4H_9-Cys^4$ OMe in the NaS-*tert*-C_4H_9/**2** = 1/1 solution is evidence for this turning of the C-terminus. The turning will enable the formation of mulitple hydrogen bonds, Cys^i S−Y H_N, Cys^i S−Cys^{i+3} H_N, Cys^i CO−Gly^{i+4} H_N, Cys^{i+3} S−Ala^{i+5} H_N, and Ala^{i+5} CO−Cys^i H_N, as observed in the cores of AT,[2] ZF motifs,[3] and *D. vulgaris* rubredoxin (Figure 1b).

Acknowledgment. This work was supported by a Grant-in-Aid for Scientific Research No. 05209227 from the Ministry of Education, Sports, Science and Culture of Japan.

Supporting Information Available: Text describing automatic simulated annealing, figures showing a ROESY spectrum, XANES spectra, structures, and histograms, and tables listing vibrational frequencies and force constants and Cartesian coordinates (40 pages). Ordering information is given on any current masthead page.

IC961009Y