Artificial Zinc Finger Peptide Containing a Novel His_4 Domain

Yuichiro Hori, Kazuo Suzuki, Yasushi Okuno, Makoto Nagaoka, Shiroh Futaki, and Yukio Sugiura*

Contribution from the Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Received November 15, 1999

Abstract: Zinc finger constitutes one of the most common DNA binding motifs. Although zinc finger proteins consisting of Cys_2His_2, Cys_3His, Cys_4, and Cys_6 domains are known in nature, a novel His_4 zinc finger protein has never been observed. Herein, we have created the first artificial His_4-type zinc finger protein (H_4Sp1) engineered by Cys → His mutations of the Cys_2His_2-type zinc finger transcription factor Sp1. The CD features of the single finger H_4Sp1f2 and three-finger H_4Sp1 clearly demonstrate the folding of the mutant His_4 peptides by complexation with Zn(II). The NMR study of Zn(II)-H_4Sp1f2 reveals that some distortions of the helical region occur due to Zn(II) coordination. The gel mobility shift assay and DNase I footprinting analysis strongly show the binding of Zn(II)-H_4Sp1 to the GC-box site of duplex DNA. The methylation interference pattern of Zn(II)-H_4Sp1 binding significantly resembles that of the corresponding C_2H_2Sp1 binding. The present artificial peptide H_4Sp1 is the first example of a zinc finger containing the His_4 domain. Of special interest is the fact that the zinc finger domains of H_4Sp1 are folded (although not identical to the native structure) and bind DNA similar to wild-type C_2H_2Sp1.

Introduction

Metalloproteins play important roles in gene regulation, and some metal ions also participate in the transcriptional process.[1−8] In particular, zinc ion is essential as a structural factor for zinc finger proteins, which constitute one of the most common DNA binding motifs.[9,10] Zinc finger proteins acquire DNA binding ability by Zn(II)-complexation.[11] On the other hand, it has been reported that other metal ions such as Co(II), Cd(II), Pb(II), Ni(II), and Fe(II) can interact with zinc finger proteins by substitution with Zn(II).[12−18] Interestingly, in some of these "substituted" zinc finger motifs, metal ions influence not only the secondary and tertiary structures of the metalloproteins but also functions such as sequence preference[16] and DNA cleavage.[17] Therefore, the utilization of metal-ligation is advantageous for designing the minidomain possessing DNA binding ability. Such a metal-containing minidomain, namely the "metallofinger motif", is expected to provide valuable information about metallopeptide−DNA interactions. In addition, artificial metallofinger motifs may be useful as a gene therapeutic agent and a tool for genetic engineering.

In nature, Cys_2His_2-, Cys_3His-, Cys_4-, and Cys_6-type zinc fingers exist. Among them, the Cys_2His_2-type zinc finger motif especially possesses the following fascinating characteristics: (1) a compact $\beta\beta\alpha$ fold is acquired by Zn(II)-coordination to bind the asymmetric DNA sequence, (2) one finger recognizes 3 to 4 base pairs by the side chains of amino acids located on the recognition helix, and (3) extended recognition can be attained by tandem repeating.[19−25] The mini $\beta\beta\alpha$ fold of the Cys_2His_2-type zinc finger motif is an attractive framework for designing a novel metallofinger motif.[9,26] The previous His → Cys conversions in the Cys_2His_2-type zinc finger peptide of Zif268 clearly demonstrated that the Cys_4 single mutation of finger 2 or double mutations of three finger domains lost DNA binding affinity.[27]

At the present stage, no newly designed metallofingers that have evident DNA recognition ability are known. In addition,

* To whom correspondence should be addressed.

(1) O'Halloran, T. V. *Science* **1993**, *261*, 715−725.
(2) Winge, D. R. *Prog. Nucleic Acids Res. Mol. Biol.* **1998**, *58*, 165−195.
(3) Escolar, L.; Pérez-Martín, J.; de Lorenzo, V. *J. Bacteriol.* **1999**, *181*, 6223−6229.
(4) Bauer, C. E.; Elsen, S.; Bird, T. H. *Annu. Rev. Microbiol.* **1999**, *53*, 495−523.
(5) Outten, C. E.; Outten, F. W.; O'Halloran, T. V. *J. Biol. Chem.* **1999**, *274*, 37517−37524.
(6) Carrion, A. M.; Link, W. A.; Ledo, F.; Mellstrom, B.; Naranjo, J. R. *Nature* **1999**, *398*, 80−84.
(7) Rutherford, J. C.; Cavet, J. S.; Robinson, N. J. *J. Biol. Chem.* **1999**, *274*, 25827−25832.
(8) Posey, J. E.; Hardham, J. M.; Norris, S. J.; Gherardini, F. C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10887−10892.
(9) Klug, A.; Shwabe, J. W. R. *FASEB J.* **1995**, *9*, 597−604.
(10) Berg, J. M.; Shi, Y. *Science* **1996**, *271*, 1081−1085.
(11) Miller, J.; Mclachlan, A. D.; Klug, A. *EMBO J.* **1985**, *4*, 1609−1614.
(12) Kuwahara, J.; Coleman, J. E. *Biochemistry* **1990**, *29*, 8627−8631.
(13) Predki, P. F.; Sarkar, B. *J. Biol. Chem.* **1992**, *267*, 5842−5846.
(14) Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937−940.
(15) Payne, J. C.; ter Horst, M. A.; Godwin, H. A. *J. Am. Chem. Soc.* **1999**, *121*, 6850−6855.
(16) Nagaoka, M.; Kuwahara, J.; Sugiura, Y. *Biochem. Biophys. Res. Commun.* **1993**, *194*, 1515−1520.
(17) Conte, D.; Narindrasorasak, S.; Sarkar, B. *J. Biol. Chem.* **1996**, *271*, 5125−5130.
(18) Krizek, B. A.; Berg, J. M. *Inorg. Chem.* **1992**, *31*, 2984−2986.
(19) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635−637.
(20) Klevit, R. E.; Herriot, J. R.; Horvath, S. J. *Proteins* **1990**, *7*, 215−226.
(21) Omichinski, J. G.; Clore, G. M.; Appella, E.; Sakaguchi, K.; Gronenborn, A. M. *Biochemistry* **1990**, *29*, 9324−9334.
(22) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809−817.
(23) Pavletich, N. P.; Pabo, C. O. *Science* **1993**, *261*, 1701−1707.
(24) Elrod-Erickson, M.; Rould, M. A.; Nekludova, L.; Pabo, C. O. *Structure* **1996**, *4*, 1171−1180.
(25) Elrod-Erickson, M.; Benson, T. E.; Pabo, C. O. *Structure* **1998**, *6*, 451−464.
(26) Imanishi, M.; Hori, Y.; Nagoka, M.; Sugiura, Y. *Eur. J. Pharm. Sci.* In press.
(27) Green, A.; Sarkar, B. *Biochem. J.* **1998**, *333*, 85−90.

10.1021/ja994009g CCC: $19.00
Published on Web 07/29/2000

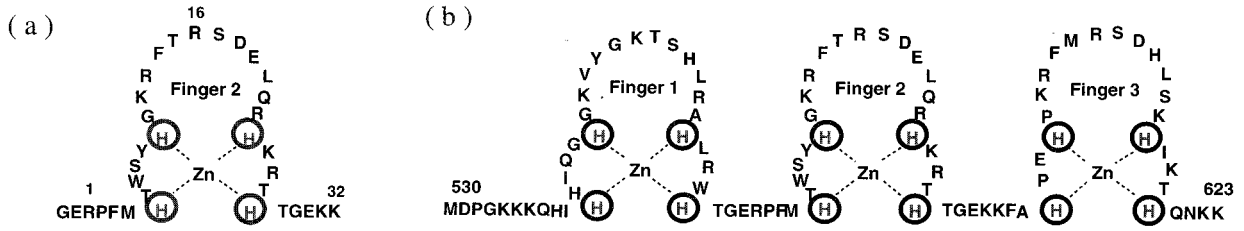

Figure 1. Novel zinc finger protein (H_4Sp1f2 (a) and H_4Sp1 (b)) containing His_4 domain.

zinc finger proteins containing the His_4 domain have never been found or engineered. Herein, we have created the first novel His_4-type zinc finger protein by Cys → His mutations of a Cys_2-His_2-type zinc finger in transcription factor Sp1 (Figure 1). Of special interest are the peptide folding and DNA recognition abilities of the artificial His_4 zinc finger protein.

Experimental Section

Chemicals. T4 polynucleotide kinase and the restriction enzymes were purchased from New England Biolabs (Beverly, MA), except for *Sal*I and *Bam*HI which were obtained from Takara Shuzo (Kyoto, Japan) and Nippon Gene (Tokyo, Japan), respectively. [γ-^{32}P]ATP was supplied by DuPont. Dr. R. Tjian kindly provided the plasmid pBS-Sp1-fl. All other chemicals were of commercial reagent grade.

Peptide Synthesis. The synthesis of the His_4 mutant peptide (H_4Sp1f2) of the middle finger of the three-finger Sp1 was conducted by Fmoc-solid-phase synthesis on a Rink amide resin. The peptide chain was constructed with a Shimadzu PSSM-8 synthesizer using its standard protocol with the benzotriazole-1-yloxytrispyrrolidinophosphonium hexafluorophosphate (PyBOP)−1-hydroxybenzotriazol(HOBt)−*N*-methylmorpholine (NMM) coupling system. The protected peptide resin of H_4Sp1f2 was treated with trifluoroacetic acid−ethanedithiol (95:5) at room temperature for 2 h, followed by HPLC-purification on a μBondesphere$_5C_4$-300 (19 × 150 mm) column. The fidelity of the product was confirmed by time-of-flight mass spectrometry (TOFMS) using a Kratos Kompact MALDI 4: calculated mass = 4224.71 g/mol and observed mass = 4225.30 g/mol.

Preparations of Zinc Finger Proteins C_2H_2Sp1 and H_4Sp1. C_2H_2-Sp1, which is the alias for Sp1(530−623), is coded on the plasmids, pUCSp1(530−623) and pEVSp1(530−623), as previously described.[28] The His_4 mutant, H_4Sp1, was constructed by mutation of six cysteine residues to histidine residues in three Sp1's zinc finger domains. A DNA fragment coding H_4Sp1 was generated from pUCSp1(530−623) by means of a PCR-based site-directed mutagenesis and the sequence was exactly confirmed by DNA sequence analysis using the BcaBEST dideoxy sequencing kit (Takara Shuzo, Kyoto). The amplified fragment was cut out and inserted into a similarly digested pEV3b.[28] Both zinc finger proteins, C_2H_2Sp1 and H_4Sp1, were overexpressed in *Escherichia coli* strain BL21(DE3)pLysS as previously described.[12] Purification procedures were performed at 4 °C. The cells of *E. coli* were resuspended and lysed in PBS buffer [130 mM NaCl, 2.7 mM KCl, and 10 mM phosphate buffer (pH 7.6)]. After centrifugation, the supernatant, which contained the soluble form of the zinc finger proteins, was purified by cation exchange chromatography (High S and UNO S-1, Bio-Rad) and gel filtration (Superdex 75, Amersham Pharmacia Biotech) with Tris buffer [10 mM Tris (pH 8.0), 50 mM NaCl, and 1 mM dithiothreitol]. To reconstitute H_4Sp1 with Zn(II), Ni(II), Cd(II), Co(II), or Cu(II), the insoluble form was also purified from the pellet after centrifugation of the lysed *E. coli* cells. The pellet was lysed in the PBS buffer containing 8 M urea and 10 mM chelating agent (EDTA or 1,10-phenanthroline), and then purified according to the same procedure as described above. Refolding was conducted by heating the purified H_4Sp1 at 65 °C for 30 min, and by cooling it gradually in 10 mM Tris buffer (pH 8.0) containing 50 mM NaCl, 1 mM dithiothreitol, and 125 μM $ZnCl_2$, $Ni(NO_3)_2$, $CdCl_2$, $Co(NO_3)_2$, or $CuSO_4$.

Circular Dichroism (CD) Measurement. The CD spectra of the synthetic single zinc finger peptides, C_2H_2Sp1f2 and H_4Sp1f2, were recorded on a Jasco J-720 spectropolarimeter in Tris buffer (pH 7.5) containing 50 mM NaCl and 5% glycerol at 4 °C. Similarly, CD measurements of the three zinc finger proteins, C_2H_2Sp1 and H_4Sp1, were also conducted.

Electron Spin Resonance (ESR) Determination. The ESR spectrum of the 1:1 H_4Sp1f2−Cu(II) complex was recorded on a JEOL JES-TE200 ESR spectrometer in Tris buffer (pH 7.5) containing 50 mM NaCl and 5% glycerol.

Nuclear Magnetic Resonance (NMR) Experiments. In the presence of 1.5 equiv of Zn(II) ion, the complex solution of H_4Sp1f2 and Zn(II) was prepared at a concentration of 5 mM in 90% H_2O/10% D_2O and D_2O (25 mM Tris-d_{11}, pH 5.7). All the NMR spectra were recorded on a JEOL Lambda-600 spectrometer. The NOESY spectra were acquired with selective water presaturation (delays alternating with mutations for tailored excitation pulse) followed by the standard NOESY pulse train at 303 K with mixing times of 100, 200, and 300 ms. The TOCSY spectra were obtained with an 80 ms MLEV-17 spinlock duration using water suppression by a gradient tailored excitation pulse at 303 K. Spectra were typically acquired with 24 scans per t1 value for 1024 t1 values, and 2048 complex points were collected in the direct dimension. The free induction decay in both dimensions was multiplied by the phase-shifted sine bell apodization function, zero-filled, and Fourier transformed to yield 2048 by 2048 matrices. Sequential resonance assignments were determined by standard TOCSY and NOESY procedures.[29]

Structure Calculation. The initial starting structure, the H_4Sp1f2 coordinates, was generated by the substitution of two Cys side chains of C_2H_2Sp1f2 PDB coordinates for the His side chains. The starting structure was subjected to a simulated annealing protocol at 600 K within XPLOR. Distance restraints were obtained from the NOESY spectra recorded with mixing times of 200 ms. The H−H distances were classified into three categories: 2.0−2.5, 2.0−3.5, and 2.0−5.0 Å, corresponding to strong, medium, and weak NOEs, respectively. Zinc−ligand bonds were assigned equilibrium distances of 2.0 Å using NOE constraints.

Gel Mobility Shift Assay. The pBS-GC, which codes the "GC box (5′-GGGGCGGGGCC-3′)" as the binding sites for Sp1, was constructed as previously described.[30] The *Hin*dIII-*Xba*I fragment (41 bp) was cut out from the pBS-GC, ^{32}P-labeled at the 5′-end with T4 polynucleotide kinase, and purified with polyacrylamide gels. Binding reaction mixtures (final volume, 20 μL) contained 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 100 μM $ZnCl_2$, 1 mM dithiothreitol, 25 ng/μL poly(dI-dC) (Amersham Pharmacia Biotech), 0.05% Nonidet P-40, 5% glycerol, the 5′-end-labeled DNA fragment (∼100 pM), and 0−4 μM zinc finger proteins purified from the soluble fractions. After incubation at 20 °C for 30 min, the sample was run on an 8% polyacrylamide nondenaturing gel with Tris-borate buffer [89 mM Tris-HCl (pH 8.0) and 89 mM boric acid]. The bands were visualized by autoradiography.

DNase I Footprinting Experiment. The DNase I footprinting analysis was performed according to the method of Brenowitz et al.[31] A substrate GC-box fragment was cut out from the pCPA5I constructed as previously described.[27] Binding reaction mixtures contained 10 mM

(28) Nagaoka, M.; Sugiura, Y. *Biochemistry* **1996**, *35*, 8761−8768.

(29) Würtrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986.

(30) Yokono, M.; Saegusa, N.; Matsushita, K.; Sugiura, Y. *Biochemistry* **1998**, *37*, 6824−6832.

(31) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132−181.

Tris buffer (pH 8.0), 50 mM NaCl, 100 μM $ZnCl_2$, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 25 ng/mL sonicated calf thymus DNA, 0.05% Nonidet P-40, 5% glycerol, the 5′-end labeled *Xba*I-*Sal*I fragment (148 bp) of pCPA5I (approximately 20 kcpm), and 0−2.4 μM zinc finger proteins. After incubation at 20 °C for 30 min, the sample was digested with DNase I (7 mU) at 20 °C for 2 min. The reaction was stopped by the addition of 20 μL of DNase I stop solution (0.1 M EDTA and 0.6 M sodium acetate) and 150 μL of ethanol. After ethanol precipitation, the cleavage products were analyzed on a 10% polyacrylamide/7 M urea sequencing gel. The bands were visualized by autoradiography.

Methylation Interference Analysis. Methylation interference experiments were performed as previously described.[32] The binding reaction mixtures contained 10 mM Tris buffer (pH 8.0), 50 mM NaCl, 100 μM $ZnCl_2$, 1 mM dithiothreitol, 25 ng/μL poly(dI-dC), 0.05% Nonidet P-40, 5% glycerol, 50 or 250 nM zinc finger proteins, and the 5′-end labeled methylated *Xba*I-*Hind*III fragment of pBS-GC (approximately 600 kcpm). Densitometric measurements were obtained using the NIH Image (version 1.61). On the basis of the cutting probabilities F and B, where F represents the intensity of the typical band in the free lane and B that in the bound lane, the extent of interference was estimated for each base as the ratio B/F. Band intensities between the free and bound lanes were corrected by comparing nucleotides outside the binding site.

Results and Discussion

CD Studies of Single Zinc Finger H_4Sp1f2. To examine the folding properties of the zinc finger domain with Cys → His mutations, the mutant peptide (H_4Sp1f2; GERPFMHTWSYHGKRFTRSDELQRHKRTHTGEKK) corresponding to the middle zinc finger (C_2H_2Sp1f2) of Sp1(530−623) was synthesized and investigated using CD measurements (Figure 2a). The CD spectrum of H_4Sp1f2 in the absence of Zn(II) showed that this peptide possesses random coil features. The addition of Zn(II) to apo-H_4Sp1f2 induced negative Cotton effects in the far-UV region with a minimum at 206 nm and a shoulder around 222 nm, indicating that H_4Sp1f2 forms an ordered secondary structure with a helical conformation by Zn(II) coordination. Indeed, the removal of Zn(II) by EDTA brought about an evident decrease in these CD bands (data not shown). These Zn(II)-dependent CD features clearly demonstrate the folding of H_4Sp1f2 by complexation with Zn(II) ion. As for the mean residue ellipticities at 222 nm, $[\theta]_{222}$, indicating the helical content, the value of Zn(II)-H_4Sp1f2 ($[\theta]_{222} = -6407$) was remarkably close to that of Zn(II)-C_2H_2Sp1f2 ($[\theta]_{222} = -6150$). The CD spectral feature of Zn(II)-H_4Sp1f2 was quite similar, but not identical to that of Zn(II)-C_2H_2Sp1f2.[33] This result suggests that the finger domain conformation of the former is not entirely the same as that of the latter.

In the presence of other metals, the CD spectra of H_4Sp1f2 were also recorded (Figure 2b). The spectra in the presence of Ni(II), Cd(II), or Co(II) gave negative Cotton effects in the far-UV region. The spectral features suggest that these metals also caused folding of H_4Sp1f2, although their spectra are characteristic of the mixture of coiled and helical structures.[37] The $[\theta]_{222}$ values in their spectra were less than that observed in the presence of Zn(II) (the $[[\theta]_{222}$ values of Zn(II)-, Ni(II)-, Cd(II)-, or Co(II)-bound H_4Sp1f2: −6407, −4231, −3995, and −3565, respectively), suggesting that the α-helix content in Ni(II)-, Cd(II)-, or Co(II)-H_4Sp1f2 is smaller than that of Zn(II)-H_4Sp1f2. The results reveal that the Zn(II) ion facilitates the formation of a more ordered conformation of H_4Sp1f2 than Ni(II), Cd(II), or Co(II) ion. On the other hand, the CD spectrum of Cu(II)-H_4Sp1 was distinct from that of Zn(II)-H_4Sp1f2 (Figure 2c). Evidently, Cu(II) induced quite a different folding of H_4Sp1f2 from that for the case of Zn(II), because negative Cotton effects typical of the α-helix were not detected. In addition, the ESR parameters ($A_{\|} = 190$ G, $g_{\|} = 2.213$, and $g_{\perp} = 2.059$) were characteristic of a distorted square-planar geometry.[38,39] Therefore, these results suggest that Zn(II) is the most effective for the formation of an ordered structure in H_4Sp1f2.

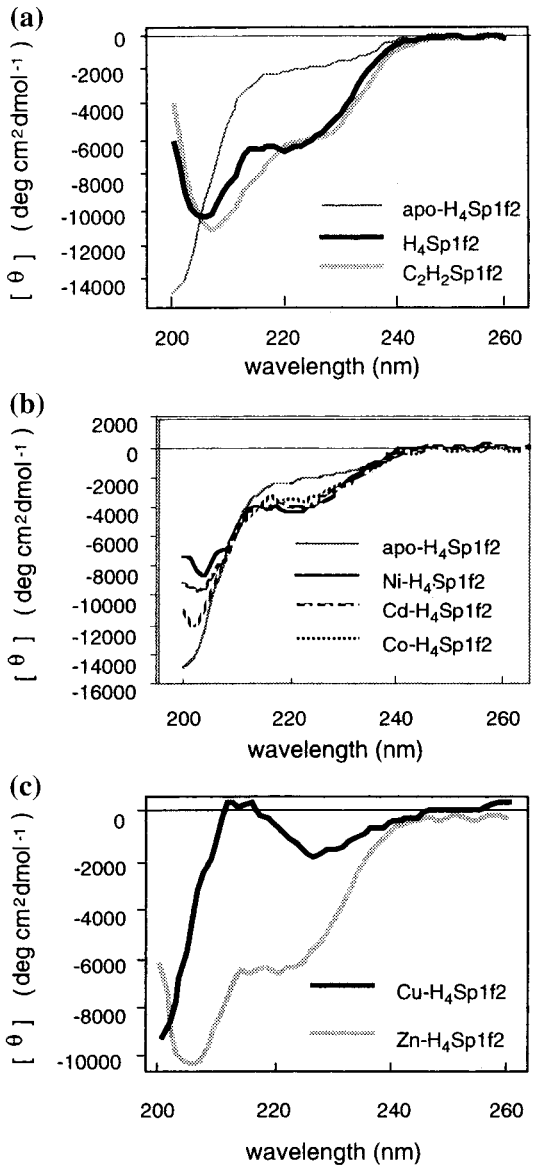

Figure 2. (a) CD spectral features of the zinc finger domain of H_4Sp1f2 (24.9 μM) and C_2H_2Sp1f2 (26.3 μM) in the absence or presence of Zn(II) (50 μM). (b) CD spectra of H_4Sp1f2 (25 μM) in the absence or presence of Ni(II) (50 μM), Cd(II) (500 μM), and Co(II) (76.1 μM). (c) CD spectra of H_4Sp1f2 (25 μM) in the presence of Zn(II) (50 μM) and Cu(II) (372 μM).

NMR Experiments of Zn(II)-H_4Sp1f2. To clarify the structural alteration of the zinc finger domain by Cys → His

(32) Kuwahara, J.; Yonezawa, A.; Futamura, M.; Sugiura, Y. *Biochemistry* **1993**, *32*, 5994−6001.

(33) In the presence of Zn(II), the CD spectrum of C_2H_2Sp1f2 showed negative ellipticities including a minimum at 208 nm and a shoulder around 225 nm characteristic of natural C_2H_2-type zinc finger peptides containing the α-helical structure.[12,34−36]

(34) Frankel, A. D.; Berg, J. M.; Pabo, C. O. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4841−4845.

(35) Parraga, G.; Horvath, S. J.; Eisen, A.; Taylor, W. E.; Hood, L. E.; Young, E. T.; Klevit, R. E. *Science* **1988**, *241*, 1489−1492.

(36) Weiss, M. A.; Mason, K. A.; Dahl, C. E.; Keutmann, H. T. *Biochemistry* **1990**, *29*, 5660−5664.

(37) Padmanabhan, S.; Baldwin, R. L. *J. Mol. Biol.* **1991**, *219*, 135−137.

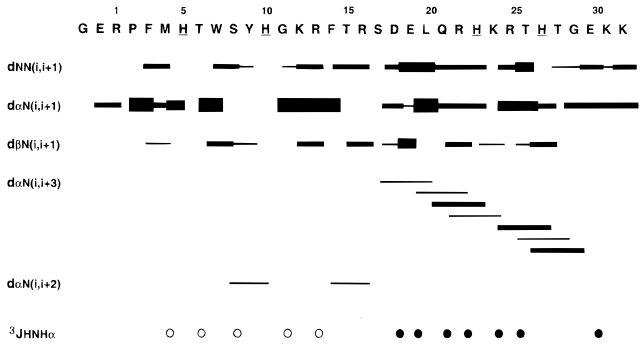

Figure 3. Summary of sequential, short-range, and medium-range NOEs observed in H_4Sp1f2. The bars below the sequences represent the obtained NOE connectivities. The thickness of the bars is a quantitative measure of the cross-peak intensity in a 200-ms NOESY spectrum. The small (<6.0 Hz) and large (>8.0 Hz) values are indicated by the filled and empty circles, respectively.

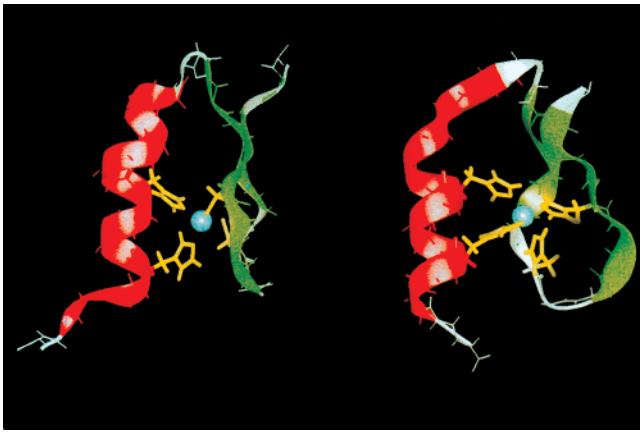

Figure 4. The proposed 3D structural models of H_4Sp1f2 (right) and the PDB structure of C_2H_2Sp1.(left).

mutations, the H_4Sp1f2-Zn(II) complex was investigated using two-dimensional NMR techniques. Sequential resonance assignments were determined by standard TOCSY and NOESY procedures. The short and medium range connective patterns of H_4Sp1f2 were considerably similar to those of C_2H_2Sp1f2[40] (Figure 3), suggesting a similarity in the secondary structures of the zinc finger domains between H_4Sp1f2 and C_2H_2Sp1f2. The amino acid residues 17−29 produced NOE patterns and $^3J_{HNH\alpha}$ coupling constants characteristic of a typical α-helix. The disappearance of $C_\alpha H_i$-NH_{i+3} NOEs between Arg22-Arg25 and His23-Thr26 indicates that the Cys → His mutations of C_2H_2Sp1f2 cause some distortions of the α-helical conformation in the neighborhood of His23 or His27. The N-terminal segments of the residues Pro2-Trp7 and Gly11-Phe14 gave large $^3J_{HNH\alpha}$ and strong $C_\alpha H_i$-NH_{i+1} connectivities characteristic of an extended (β-like) strand conformation. The $C_\alpha H_i$-NH_{i+2} NOE peak between Ser8 and His10 is also consistent with a turn in the Trp7-His10 region connecting the two extended strands. On the basis of molecular dynamics simulation from the NMR distance and dihedral restraints, a three-dimensional structure of Zn(II)-H_4Sp1f2 was generated from the PDB coordinates of the corresponding C_2H_2Sp1f2 (Figure 4).[40] The obtained model was fully consistent with the observed NMR-derived data, though the restraints of the zinc ligand bonds were essential

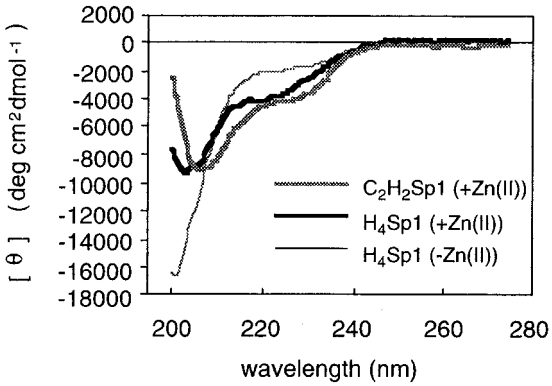

Figure 5. CD spectra of H_4Sp1 and C_2H_2Sp1 in the presence of Zn(II). The concentrations of proteins and Zn(II) ion are 9.2 and 158 μM, respectively.

for definition of the structure. In contrast to C_2H_2Sp1f2, a 3D structural model of H_4Sp1f2 showed the twisting of an extended strand in the Pro2-Trp7 region and helical unwinding around Asp18-Gln21, as indicated by the disappearance of the $C_\alpha H_i$-NH_{i+3} NOEs between them. The twisting detected in the Pro2-Trp7 region probably occurs to accommodate a tetrahedral binding geometry enforced by zinc coordination constraints.

CD Features of Three Zinc Finger H_4Sp1. To examine whether the folding of H_4Sp1 is Zn(II)-dependent as well as that of H_4Sp1f2, CD studies were carried out. The mutant protein H_4Sp1 was constructed from Sp1(530−623) (C_2H_2Sp1) by Cys → His mutations of all three of the zinc finger domains. The CD spectrum of H_4Sp1 in the Zn(II)-free buffer was characteristic of the random coil (Figure 5). In the presence of Zn(II), the negative CD signal bands near 203 and 222 nm clearly increased, demonstrating the folding of H_4Sp1 by Zn(II)-complexation. The CD feature of Zn(II)-H_4Sp1 was different from that (207 and 226 nm) of Zn(II)-C_2H_2Sp1 to some degree, suggesting that the zinc finger domain of H_4Sp1 did not form the structure identical to that of the wild-type C_2H_2Sp1. Furthermore, thermal stability of Zn(II)-H_4Sp1 was investigated by monitoring the CD signal change and then compared with that of Zn(II)-C_2H_2Sp1. Figure 6 shows CD spectra of C_2H_2-Sp1 (a) and H_4Sp1 (b) at 20−80 °C in the presence of Zn(II). The conformation of Zn(II)-H_4Sp1 was significantly unfolded at 80 °C, although Zn(II)-C_2H_2Sp1 maintained its secondary structures even at that temperature. The decreased thermal stability of the His_4 mutant may be due to weakened hydrophobic interaction by an extended β-like strand conformation of the N-terminal segment and a distorted α-helical conformation near His23 or His27.

DNA Binding Abilities of Zn(II)-H_4Sp1. The results of the gel mobility shift assays clearly showed that Zn(II)-H_4Sp1 was bound to the DNA fragment (41 bp) containing an Sp1 recognition site GC box (5′-GGGGCGGGGCC-3′) (Figure 7). From the evidence for the monomeric binding of the C_2H_2-type zinc finger to the single binding site,[22,23] the binding mode of Zn(II)-H_4Sp1 is also monomeric because the mobility of the shifted band was the same as that of Zn(II)-C_2H_2Sp1 (data not shown). However, the DNA binding affinity (Kd = 85 nM) of H_4Sp1 for the GC box was lower than that (Kd = 3.5 nM)[30] of C_2H_2Sp1. Furthermore, the Zn(II)-dependent DNA binding ability of H_4Sp1 was shown by the decrease in the DNA binding by EDTA (data not shown). Figure 8 displays the DNase I footprinting pattern of Zn(II)-H_4Sp1 for a 148-bp DNA fragment containing the GC box. In the presence of Zn(II), H_4Sp1 almost protected the residues of the GC box at 2.4 μM from DNase I cleavage. In addition, the binding of Zn(II)-H_4Sp1 to the GC

(38) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. *Chem. Rev.* **1992**, *92*, 521−524.
(39) Peisach, J.; Blumberg, W. E. *Arch. Biochem. Biophys.* **1974**, *165*, 691−708.
(40) Narayan, V. A.; Kriwacki, R. W.; Cardonna, J. P. *J. Biol. Chem.* **1997**, *272*, 7801−7809.

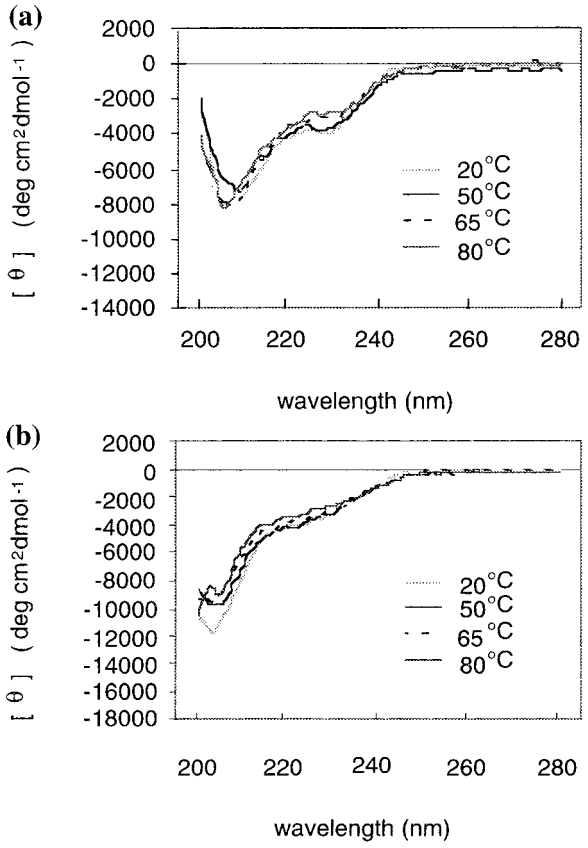

Figure 6. CD spectra of C_2H_2Sp1 (a) and H_4Sp1 (b) at 20, 50, 65, and 80 °C in the presence of Zn(II).

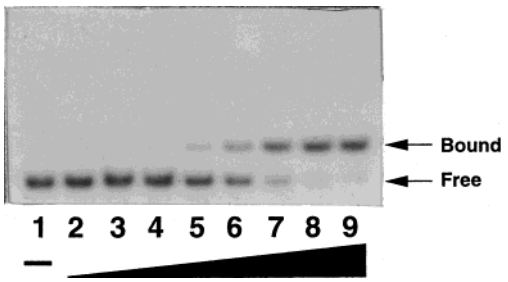

Figure 7. Gel mobility shift assay for H_4Sp1 binding to GC-box DNA (41 bp) in the presence of Zn(II) (100 μM). Lanes 1−9 represent 0, 3.9, 7.8, 15.6, 31.3, 62.5, 125, 250, and 500 nM of H_4Sp1, respectively.

box caused a hypersensitive cleavage at A(13) outside the GC box. This hypersensitive cutting has also been detected in the case of the corresponding C_2H_2Sp1.[30,41] These results strongly indicate that the H_4Sp1 still retained the DNA binding ability despite the Cys → His mutation.

Previous studies have demonstrated that C_2H_2Sp1 can bind DNA by interacting with Co(II), Cd(II), and Ni(II) as well as Zn(II).[12,13] To clarify whether the present H_4Sp1 gains the DNA binding ability by various metal complexations, H_4Sp1 was reconstructed with Zn(II), Ni(II), Cu(II), Co(II), or Cd(II) from insoluble fractions during the purification step, and a gel mobility shift assay was conducted. While Zn(II)-reconstituted H_4Sp1 was really bound to a GC-box fragment, under this experimental condition, the DNA binding activity was not detected in the case of Ni(II)-, Cu(II)-, Co(II)-, and Cd(II)-reconstituted H_4Sp1 (data not shown). From the CD studies of the H_4Sp1f2 described above, only the Zn(II) binding to its His_4

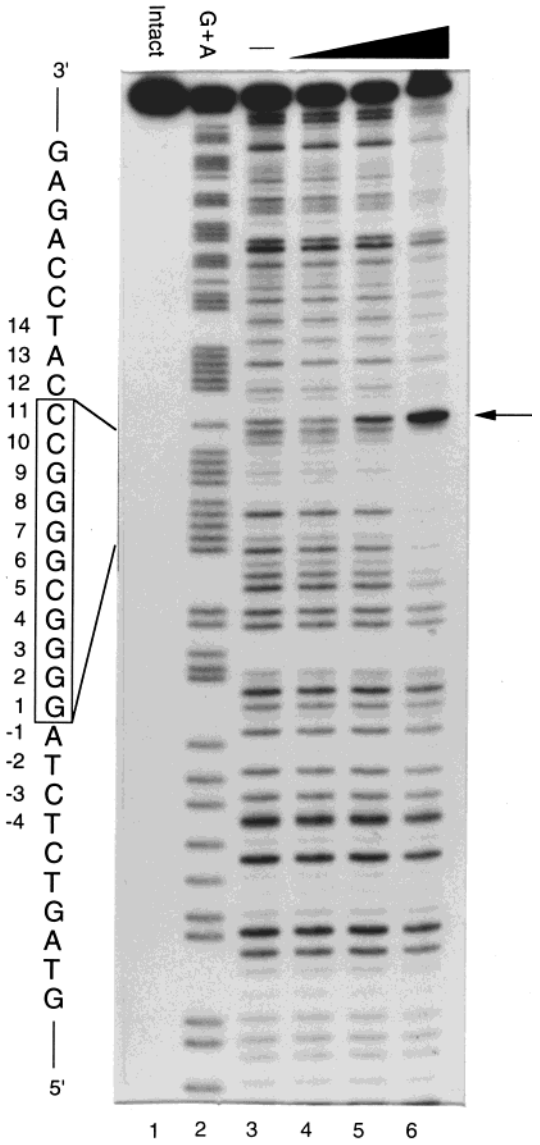

Figure 8. DNase I footprinting analysis of H_4Sp1 to GC box DNA in the presence of Zn(II) (100 μM). Lanes 1 and 2 represent intact GC box DNA and G+A of the Maxam−Gilbert reaction, respectively. Lanes 3−6 indicate the samples containing peptide concentrations of 0, 0.5, 1.2, and 2.4 μM, respectively.

domain highly induced an ordered conformation with the secondary structure. Therefore, these results indicate that effective folding by Zn(II) binding is essential for the DNA binding ability of H_4Sp1.

DNA Base Recognitions of Zn(II)-H_4Sp1 Binding to the GC Box. When Zn(II)-H_4Sp1 was bound to the GC-box fragment (41 bp), its methylation interference pattern was compared with that of Zn(II)-C_2H_2Sp1 (Figure 9). In the case of H_4Sp1, strong base contacts with G(2), G(3), G(4), and G(6) in the guanine-rich strand (G-strand) and G(5′) in the cytosine-rich strand (C-strand) were detected and also weak contacts with G(1) and G(7) in the G-strand and G(11′) in the C-strand were observed. The interference feature remarkably resembled that of the wild-type C_2H_2Sp1, indicating similar DNA binding modes between H_4Sp1 and C_2H_2Sp1.

In the Cys_4 mutants of Zif268 obtained by the His → Cys conversion, each mutation in finger 1 or 3 of the three finger domains preserved the DNA recognition ability.[27] However, the mutation in finger 2 or the double mutations of the three finger domains lost the DNA binding affinity. Presumably, structural

(41) Kamiuchi, T.; Abe, E.; Imanishi, M.; Kaji, T.; Nagaoka, M.; Sugiura, Y. *Biochemistry* **1998**, *37*, 13827−13834.

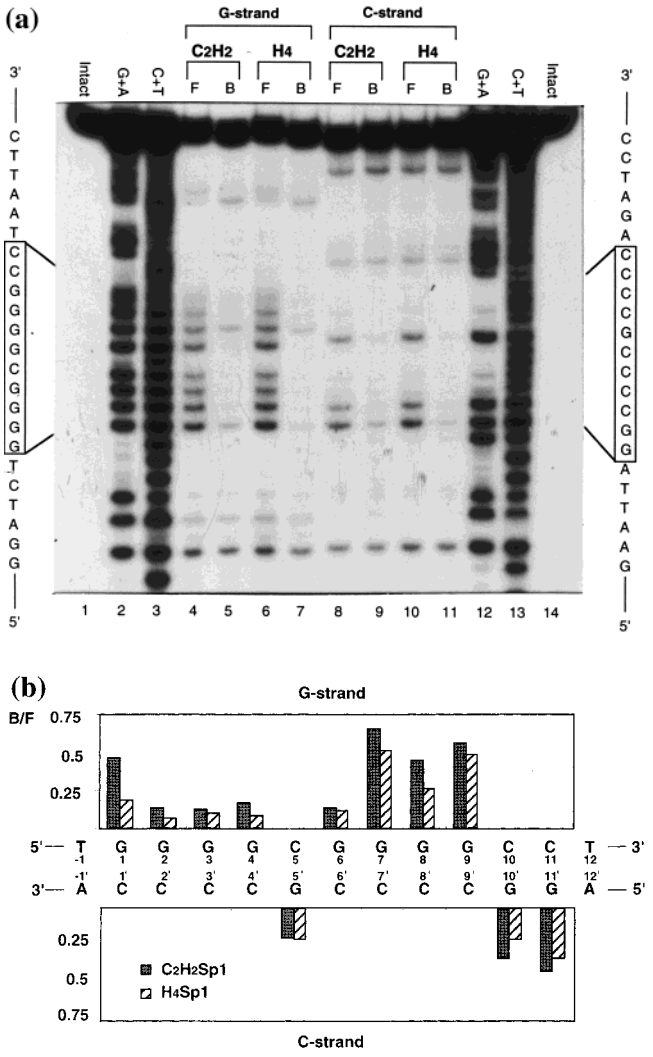

Figure 9. (a) Methylation interference analyses of C_2H_2Sp1 (lanes 4, 5, 8, and 9) and H_4Sp1 (lanes 6, 7, and 11) bindings to the G-strand (lanes 4−7) and the C-strand (lanes 8−11) of GC box DNA in the presence of Zn(II) (100 μM). Lanes 4, 6, 8, and 10 show protein-free (F) DNA samples, and lanes 5, 7, 9, and 11 peptide (C_2H_2Sp1: 50 nM and H_4Sp1: 250 nM)-bound (B) DNA samples. Lanes 1 and 14 present intact DNA, lanes 2 and 12, the G+A of the Maxam−Gilbert reaction, and lanes 3 and 13, the C+T reaction. (b) Histogram showing the extent of methylation interference by H_4Sp1 and C_2H_2Sp1. The extent of the interference was estimated with a densitometer and also calculated as the ratio of cutting probabilities for the bands (B/F).

distortion around the metal center by His → Cys mutations induces the loss of DNA binding ability. Although the Cys → His mutation results in a decreased DNA binding affinity, on the other hand, the present peptide H_4Sp1 maintains the DNA binding ability. Despite the structural alteration or distortion of the zinc finger domain as clearly indicated by the CD and NMR evidence of H_4Sp1f2, it is of special interest that specific interactions between the recognition helix of the finger 2 and its subsite, 5′-GCG-3′, are retained.[42] In addition, fingers 1 and 3 of H_4Sp1 also preserve the DNA recognition ability of the wild-type C_2H_2Sp1. Several reasons are considered for the fact that H_4Sp1 and C_2H_2Sp1 show similar DNA recognition, despite the somewhat different helical conformation between their zinc finger domains. First, the DNA binding to Zn(II)-H_4Sp1 may cause an additional structural change and result in a specific DNA recognition analogous to Zn(II)-C_2H_2Sp1. Indeed, it is known that the specific DNA binding induces a structural alteration in the basic region of GCN4.[43−45] Second, Zn(II)-H_4Sp1 may recognize DNA bases in a different manner from Zn(II)-C_2H_2Sp1. Namely, the mutant peptide H_4Sp1 compensates for specific contacts with particular bases by a directional change in the side chain that is essential for DNA recognition. As a result, the recognition helix of Zn(II)-H_4Sp1 may be located in a different orientation to DNA from that of Zn(II)-C_2H_2Sp1. To expand the discussion of DNA-H_4Sp1 contact on the basis of the NMR structure, we tentatively superimposed the structures of H_4Sp1f2 and C_2H_2Sp1f2 on Zif268 finger 1 in the Zif268-DNA crystal structure. It has been reported that overlay of C_2H_2-Sp1f2 with Zif268 finger 1 in the cocrystal structure almost exactly overlaps, and that Arg16 and Arg22 residues of C_2H_2-Sp1f2 are able to interact with the underlined guanine bases of the GCG subsite at the major groove in an orientation similar to that observed for Zif268 finger 1.[40] The present superimposition suggests that the orientation and docking mode of H_4Sp1f2 in the DNA major groove are considerably similar to those observed for Zif268 finger 1 and C_2H_2Sp1f2. In particular, the α-helix region of the residues 19−29 is remarkably superposed between H_4SP1f2 and C_2H_2SP1f2, although the β-sheet portion somewhat deviates. As demonstrated in the Zif268−DNA complex, the Lys12 and His23 residues seem to serve to the positioning of the H_4Sp1f2 zinc finger by interaction with DNA backbone phosphates. Indeed, these Lys and His residues are conserved in the zinc finger domains of Zif268 and Sp1. The H_4 mutant retains the general $\beta\beta\alpha$ structural features of C_2H_2-wild-type Sp1 despite certain conformational differences as detected by the NMR study. As a result, Zn(II)-H_4Sp1 shows sequence specificity similar to Zn(II)-C_2H_2Sp1. In the present H_4 mutant, the lowering of DNA fitting induced by some structural distortions may lead to the decreased DNA binding affinity.

In conclusion, this paper describes the preparation, structure, and DNA binding properties of an engineered His_4 mutant of the zinc finger protein Sp1. We present the NMR structure for a single mutated zinc finger (H_4Sp1f2) from Sp1 and the DNA binding data for a three-domain mutant (H_4Sp1). The CD characteristics of H_4Sp1f2 and H_4Sp1 are also compared with those of C_2H_2Sp1f2 and C_2H_2Sp1. Interestingly, the His_4 domain is folded and recognizes the GC-box DNA. The development of such a novel peptide may also lead to a new zinc finger type or other metal finger domain. Additionally, the present results suggest that a His_4-type zinc finger protein may also exist in nature.

Acknowledgment. This study was supported in part by a Grant-in-Aid for Priority Project "Biometals" (08249103) and Scientific Research (B) (10470493) from the Ministry of Education, Science, Sports, and Culture, Japan. We are also thankful for the grant from the Houansha Foundation.

JA994009G

(42) In previous reports, the fingers 1, 2, and 3 of C_2H_2Sp1 respectively recognize their subsites, 5′-GGGCC-3′, 5′-GCG-3′, and 5′-GGG-3′ within the GC box in a overlapping manner.[30]

(43) Weiss, M. A.; Ellenberger, T.; Wobbe, C. R.; Lee, J. P.; Harrison, S. C.; Struhl, K. *Nature* **1990**, *347*, 575−578.

(44) Talanian, R. V.; McKnight, C. J.; Kim P. S. *Science* **1990**, *249*, 769−771.

(45) Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223−37.