α-Helical Linker of an Artificial 6-Zinc Finger Peptide Contributes to Selective DNA Binding to a Discontinuous Recognition Sequence†

Wei Yan,‡ Miki Imanishi,*,‡,§ Shiroh Futaki,‡ and Yukio Sugiura*,||

Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan, PRESTO, JST, Sanbancho Building, 5-Sanbancho, Chiyodaku, Tokyo 102-0075, Japan, and Faculty of Pharmaceutical Sciences, Doshisha Women's University, Kyotanabe, Kyoto 610-0395, Japan

Received April 4, 2007; Revised Manuscript Received May 21, 2007

ABSTRACT: Although many zinc finger motifs have been developed to recognize specific DNA triplets, a rational way to selectively skip a particular non-recognized gap in the DNA sequence has never been established. We have now created a 6-zinc finger peptide with an α-helix linker, Sp1ZF6$(EAAAR)_4$, which selectively binds to the discontinuous recognition sites in the same phase (10 bp gap) against the opposite phase (5 bp gap) of the DNA helix. The linker peptide $(EAAAR)_4$ forms an α-helix structure stabilized by salt bridges, and the helical length is estimated to be about 30 Å, corresponding to that of the 10 bp DNA. The gel shift assays demonstrate that Sp1ZF6$(EAAAR)_4$ preferably binds to the 10 bp-gapped target rather than the 5 bp-gapped target. The CD spectra show that the α-helical content of the $(EAAAR)_4$ linker is higher in the complex with the 10 bp-gapped target than in the complex with the 5 bp-gapped target. The present results indicate that the α-helical linker is suitable for binding to the recognition sites in the same phase and that the linker induces the loss of binding affinity to the recognition sites with the opposite phase. The engineering of a helix-structured linker in the 6-zinc finger peptides should be one of the most promising approaches for selectively targeting discontinuous recognition sites depending on their phase situations.

Control of gene expression is an important challenge in the post-genomic era, and hence, artificial DNA binding peptides with the desired sequence specificity and selectivity are required. A C_2H_2 zinc finger motif recognizes the target DNA triplet with high specificity (*1*). By connecting multiple zinc finger motifs, in tandem, artificial zinc finger peptides that regulate a specific gene can be prepared (*2*−*10*). To date, two kinds of approaches have been used to create artificial zinc finger peptides: (i) engineering key amino acid residues in base recognition to create unnatural zinc finger peptides with a novel DNA binding specificity (*11*−*17*) and (ii) connecting multiple zinc finger motifs to enlarge the DNA binding region and to elevate the sequence selectivity (*18*−*25*). In order to target a unique site in the huge genome (3×10^9 bp[1]), the length of the target sequence should be greater than 16 bp ($4^{16} = 4 \times 10^9$). Therefore, many 6-zinc finger peptides targeting the continuous 18 bp have been created by connecting six independent zinc finger motifs or two naturally occurring 3-zinc finger domains with the conserved linker (TGEKP) (*18*−*20*). However, all of the DNA triplets are still not always recognized by the engineered zinc finger motifs (*15*−*17*). One approach to solve this non-recognition problem is to design zinc finger peptides that skip non-recognized bases.

So far, we have connected two 3-zinc finger domains of transcription factor Sp1 with flexible glycine linkers (*22*). The 3-zinc finger domain of Sp1 binds to a GC-rich sequence called the GC-box (*26*−*28*). We showed that the 6-zinc finger peptides could bind to DNA with two discontinuous GC-box sequences as well as two continuous GC-box sequences (*22*, *25*). We also used eight arginine residues as the linker of two 3-zinc finger domains of Sp1 (*25*). Interestingly, the peptide with the arginine linker preferably bound to the two discontinuous GC-box sequences against the two continuous GC-box sequences. Pabo and co-workers demonstrated binding to the discontinuous recognition sites using artificial dimeric zinc finger peptides (*29*, *30*). Moore et al. designed 6-zinc finger peptides by connecting two 3-zinc finger domains with a non-sequence-specific zinc finger as the linker (*23*). They showed that the 6-zinc finger peptides with such structured linkers induce increases in the binding affinity to the discontinuous recognition sites compared to that with a flexible $TG(GSG)_5ERP$ linker. However, it is still challenging to design zinc finger peptides that exhibit selectivity for discontinuous DNA sequences with a particular length of a non-recognized gap. In addition, the relative position, namely, phase, of protein binding sites in DNA is important for gene regulation as shown by the crucial effect of DNA bending and looping on transcriptional control (*31*). DNA bending and looping bring two discrete transcrip-

† This study was supported in part by Grants-in-Aid for Scientific Research (18710183 to M. I. and 17390028 to Y. S.) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

* To whom correspondence should be addressed. Tel: +81-774-38-3212. Fax: +81-774-32-3038. E-mail: imiki@scl.kyoto-u.ac.jp (M.I.). Tel: +81-774-65-8649. Fax: +81-774-65-8652. E-mail: ysugiura@dwc.doshisha.ac.jp (Y.S.).

‡ Kyoto University.

§ PRESTO.

|| Doshisha Women's University.

[1] Abbreviations: CD, circular dichroism; Tris, tris(hydroxymethyl)-aminomethane; TN, Tris-NaCl; bp, base pair(s); a.a., amino acid(s).

10.1021/bi7006417 CCC: $37.00
Published on Web 06/29/2007

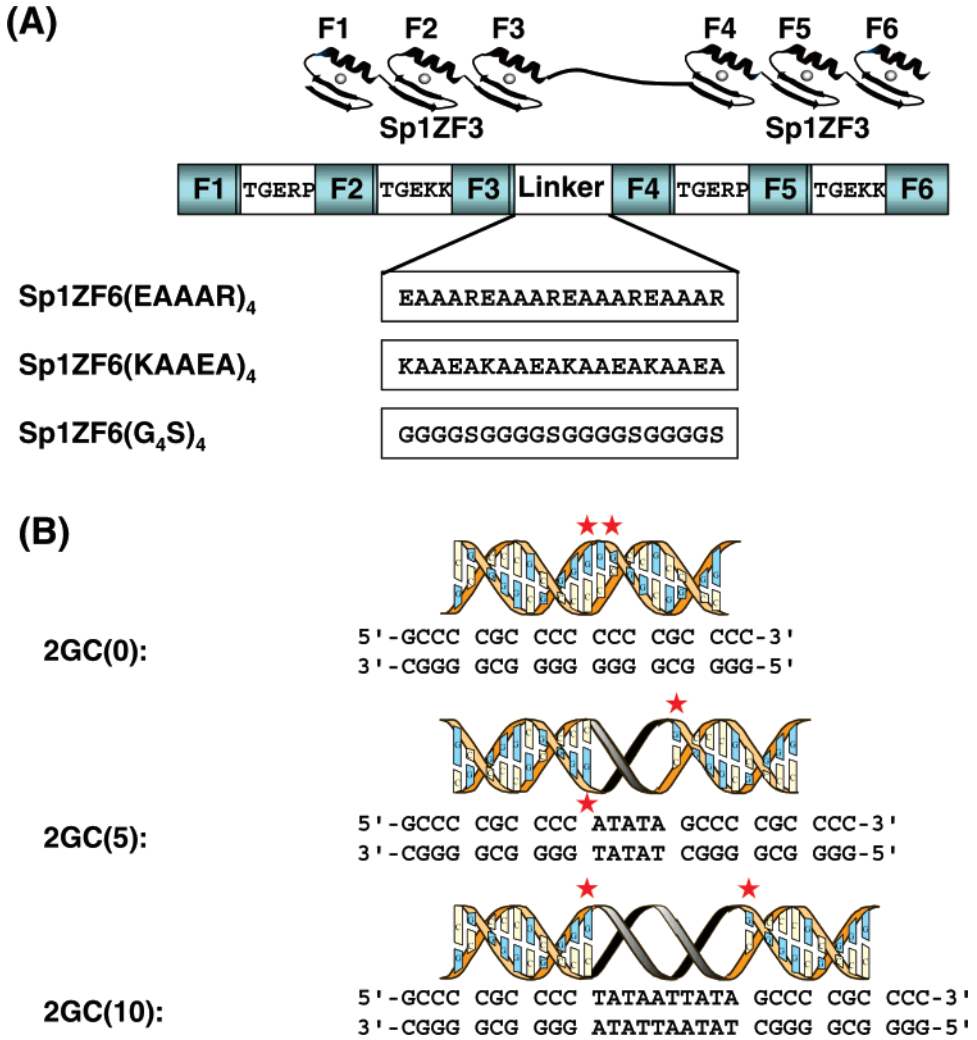

FIGURE 1: Schematic representation of Sp1ZF6(EAAAR)$_4$, Sp1ZF6(KAAEA)$_4$, and Sp1ZF6(G$_4$S)$_4$ (A) and their target DNA sequences, 2GC(0), 2GC(5), and 2GC(10) (B). The 2GC(0) sequence contains two continuous GC-box (5′-GGGGCGGGG-3′) sequences. The 2GC(5) and the 2GC(10) sequences contain two discontinuous GC-box sequences with nonbound gaps of 5 and 10 bp, respectively. The red stars indicate the last and first bases recognized by finger 3 and finger 4, respectively.

tion factor-binding sites close to each other spatially. A previous report revealed that the insertion of one helical turn of DNA to a promoter region results in a restoration of the functionality while insertion of half a helical turn leads to a loss of the functionality (*32*). Here, we focused on creating a 6-zinc finger peptide that distinguishes the discontinuous recognition sites in the same phase with a non-recognized 10 bp gap from the sites in the opposite phase with a 5 bp gap (Figure 1).

The peptides (EAAAR)$_n$ ($n > 2$) were demonstrated as model peptides with an enhanced helicity compared to the many designed α-helical models (*33*). Making use of a compact, stable, and directional α-helical structure, the peptides (EAAAR)$_n$ were applied in high-throughput screening of rapamycin-mediating FKBP12-FRB interactions (*34*) and in designing a photoswitchable peptide (*35*). In this study, we created a 6-zinc finger peptide, Sp1ZF6(EAAAR)$_4$, in which two DNA binding domains of transcription factor Sp1 are connected by an α-helical peptide sequence, (EAAAR)$_4$. For the target with a 5 bp gap, the linker between fingers 3 and 4 should be bent to cross over the DNA major groove. However, the linker can be straight when the 6-zinc finger peptide binds to the target with a 10 bp gap (Figure 1B). The length of the (EAAAR)$_4$ helix is estimated to be about 30 Å, corresponding to that of the 10 bp DNA. We expected that a 6-zinc finger peptide with an α-helix linker would have DNA binding preferences dependent on the relative positions of the discontinuous recognition sites, especially in the same phase (10 bp gap) rather than the opposite phase (5 bp gap) of the DNA helix. The DNA binding properties of Sp1ZF6(EAAAR)$_4$ for the discontinuous recognition sites with non-recognized gaps of 0, 5, and 10 bp were examined and compared to those of the 6-zinc finger peptides, Sp1ZF6(KAAEA)$_4$ with a low helical (KAAEA)$_4$ linker (*36*) and Sp1ZF6(G$_4$S)$_4$ with a flexible (G$_4$S)$_4$ linker.

MATERIALS AND METHODS

Chemicals. The T4 polynucleotide kinase and restriction enzymes were purchased from New England Biolabs. The labeled [γ-^{32}P] ATP compound was supplied by Perkin-Elmer. The plasmid pBS-Sp1-fl was kindly provided by Dr. R. Tjian. The oligonucleotides were purchased from Sigma-Aldrich and Invitrogen. All other chemicals were of commercial reagent grade.

Preparations of 6-Zinc Finger Peptides and Target DNA Fragments. Sp1ZF6$(EAAAR)_4$, Sp1ZF6$(KAAEA)_4$, and Sp1ZF6$(G_4S)_4$ were constructed as previously described by exchanging the TGEKP linker region of Sp1ZF6 (*20*) with the $(EAAAR)_4$, $(KAAEA)_4$, and $(G_4S)_4$ linkers, respectively. Sp1ZF6$(EAAAR)_4$, Sp1ZF6$(KAAEA)_4$, and Sp1ZF6$(G_4S)_4$ were overexpressed as a soluble form in *Escherichia coli* BL21(DE3)pLysS at 20 °C and purified according to the following procedure at 4 °C. *E. coli* cells were resuspended and lysed in PBS buffer. The soluble form was purified by cation-exchange chromatography using a High S Cartridge (Bio Rad) followed by a Mono S 5/50 column (Amersham). Final purification was achieved by a gel filtration technique (Superdex 75; Amersham) using TN buffer (10 mM Tris-HCl (pH 8.0), 50 mM NaCl, and 1 mM dithiothreitol). The target double stranded DNA fragments, [2GC(0)], [2GC(5)], and [2GC(10)], were prepared by annealing the 37-, 43-, and 48-base oligonucleotides and their complement oligonucleotides containing the 2GC(0): 5′-GGGGCGGGGGGG-GCGGGGC-3′, the 2GC(5): 5′-GGGGCGGGGCTATATG-GGGCGGGGC-3′, and the 2GC(10): 5′-GGGGCGGGGCT-ATAATTATAGGGGCGGGGC-3′ sequences, respectively (Figure 1).

CD Measurements. The CD spectra for all of the peptides, target DNA fragments, and the complexes of the peptide-target DNA were recorded by a Jasco J-720 spectropolarimeter using a 0.10-cm path length quartz cell at 20 °C. The CD spectra were averages of 3 or 5 scans, collected at 0.5 nm intervals between 200 and 320 nm. All of the samples were diluted to obtain 4.5 μM peptides, target DNA, or their mixtures in 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.005% Nonidet P-40, and 1 mM dithiothreitol. Before the CD measurements, the mixtures of the peptide and the target DNA fragment were incubated at 20 °C for 1 h.

Each 6-zinc finger peptide consists of two common 3-zinc finger domains of Sp1 and a disparate linker. The 3-zinc finger domain of Sp1 is well known for forming stable ββα-structures independent of the linker sequence. Provided that the contributions of the two 3-zinc finger domains and DNA in the peptide−DNA complexes are equal among the three 6-zinc finger peptides, differences in the CD spectra can be attributed to the linker region between fingers 3 and 4. Differences in the CD spectra were calculated by setting the CD spectra of Sp1ZF6$(G_4S)_4$ with flexible linker as the standards. According to eq 1, we calculated the differences in the mean residue molecular ellipticities at 222 nm, $\Delta[\theta]_{222(\text{linker})}$ (deg·cm^2·dmol^{-1}), between Sp1ZF6$(EAAAR)_4$ or Sp1ZF6$(KAAEA)_4$ and Sp1ZF6$(G_4S)_4$.

$$\Delta[\theta]_{222(\text{linker})} = (\theta_{222(\text{G4S})4} - \theta_{222(\text{linker})})/n_{(\text{linker})}Cl \quad (1)$$

In eq 1, $\theta_{222(\text{G4S})4}$ (mdeg) is ellipticity at 222 nm of free or peptide−DNA complexes of Sp1ZF6$(G_4S)_4$, $\theta_{222(\text{linker})}$ (mdeg) is ellipticity at 222 nm of free or peptide−DNA complexes of Sp1ZF6$(EAAAR)_4$ or Sp1ZF6$(KAAEA)_4$, $n_{(\text{linker})}$ is the numbers of amino acid residues in the linker between fingers 3 and 4, C (M) is the molar concentration of peptides or peptide−DNA complexes, and l (cm) is the path length of the quartz cell.

Gel Mobility Shift Assays. Gel mobility shift assays were carried out under the following conditions. Each reaction mixture contained 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM dithiothreitol, 0.005% Nonidet P-40, 5% glycerol, 5 mM $MgCl_2$, 25 ng/μL poly(dI-dC) (GE Healthcare Bio-Sciences), the 5′-end-^{32}P-labeled target DNA fragments (<20 pM), and different concentrations of the zinc finger peptides. After incubation at 20 °C for 1 h, the sample solutions were electrophoresed on an 8% nondenaturing polyacrylamide gel with Tris-borate buffer (88 mM Tris-HCl (pH 8.0) and 88 mM boric acid) at room temperature. The bands were visualized using a STORM instrument (GE Healthcare Bio-Sciences) and analyzed by ImageQuant software (Molecular Dynamics). The equilibrium dissociation constants (K_d) of each peptide−DNA complex were evaluated by fitting the experimentally obtained values of θ_b (the fraction of labeled DNA bound to the peptide) to the binding isotherm eq 2 using the Kaleida Graph program (Abelbeck software).

$$\theta_b = [\text{peptide}]/([\text{peptide}] + K_d) \quad (2)$$

Competition Experiments. The competitive DNA binding experiments were performed under the same conditions as the gel mobility shift assay, except that the non-labeled target DNA was also added as a competitor in the reaction buffer besides the labeled one. The concentrations of the peptides were determined to be in the range in which 50−70% of the labeled target DNA should be shifted. The concentrations of the non-labeled target DNA were set to be equal to those of the peptides. That is, when the ^{32}P-labeled [2GC(5)] was used, the concentration of the peptides and non-labeled target DNA was 25 nM, and when ^{32}P-labeled [2GC(10)] was used, the concentration of the peptides and non-labeled target DNA was 6 nM.

RESULTS

Creation of a 6-Zinc Finger Peptide with $(EAAAR)_4$ Linker and Target DNA. A novel 6-zinc finger peptide, Sp1ZF6-$(EAAAR)_4$, was created by linking two 3-zinc fingers of the transcription factor Sp1 with 20 amino acid residues, $(EAAAR)_4$. Sp1ZF6$(KAAEA)_4$ and Sp1ZF6$(G_4S)_4$ were also prepared as controls (Figure 1A). The DNA binding domain of the transcription factor Sp1 strongly binds to the dsDNA containing the GC-rich sequence (5′-GGGGCGGGGC-3′) called the GC-box (*26*−*28*). As the target DNA of the artificial 6-zinc finger peptides, we used [2GC(0)], [2GC-(5)], and [2GC(10)], containing two GC-box sites with a non-recognized gap of 0, 5, and 10 bp, respectively (Figure 1B).

CD Spectra of 6-Zinc Finger Peptides and Peptide−DNA Complexes. The CD spectra of the peptides at 20 °C are presented in Figure 2A. As described in the Materials and Methods section, each 6-zinc finger peptide consists of two common 3-zinc finger domains of Sp1 and a disparate linker. The 3-zinc finger domain of Sp1 is well known for forming similar ββα-structures independent of the linker sequence. Therefore, we assumed that the differences in the CD spectra of the peptides are predominantly attributed to the linker region. Sp1ZF6$(EAAAR)_4$, Sp1ZF6$(KAAEA)_4$, and Sp1ZF6-$(G_4S)_4$ showed the CD features typical of the Sp1 zinc finger peptides with negative Cotton effects in the far-UV region with a minimum around 208 nm and a shoulder between 220 and 230 nm (*37*, *38*). The negative Cotton effects at 222 nm increased in the order of Sp1ZF6$(G_4S)_4$, Sp1ZF6-$(KAAEA)_4$, and Sp1ZF6$(EAAAR)_4$. In addition, the minimum around 208 nm shifted to a longer wavelength in

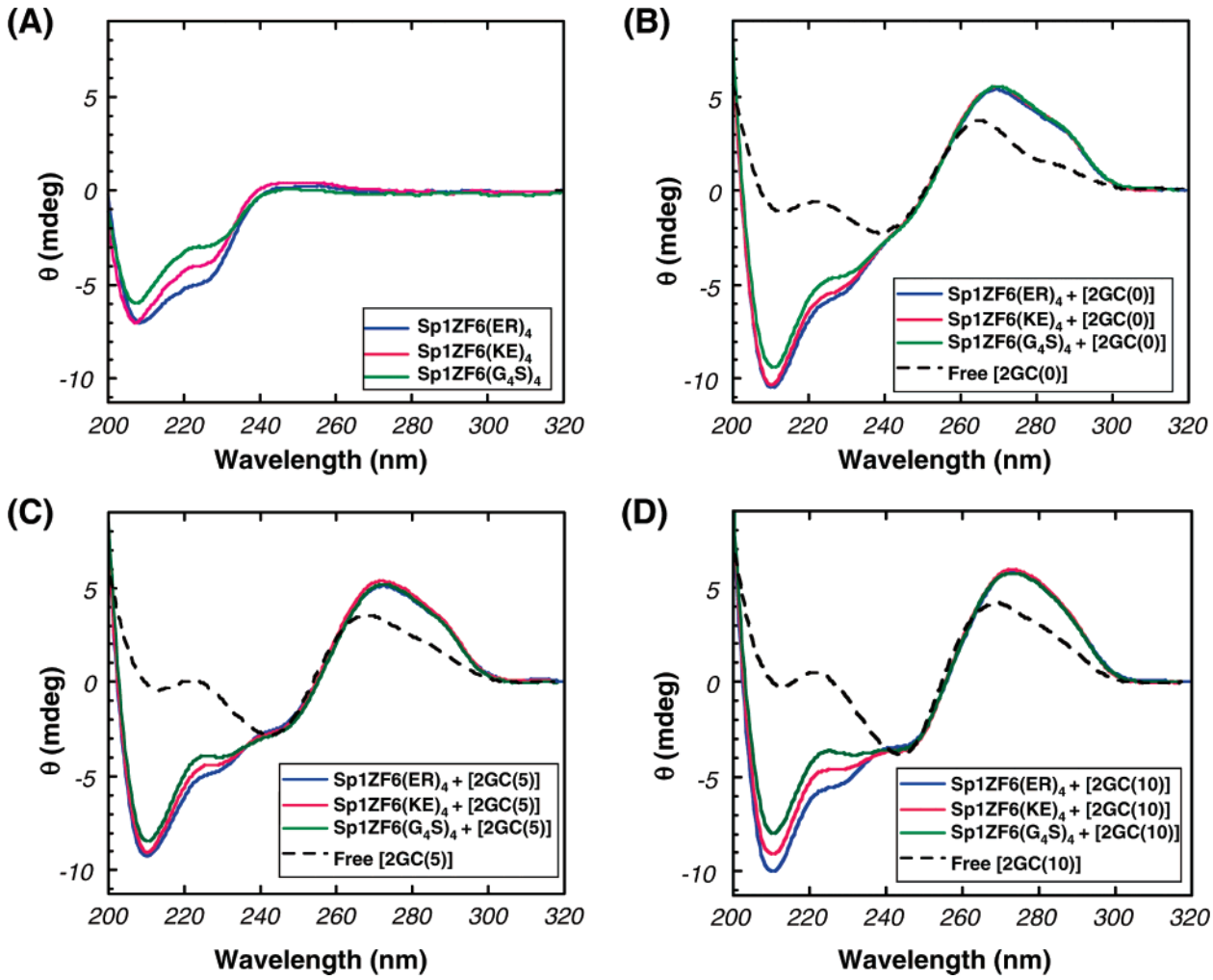

FIGURE 2: CD spectra of Sp1ZF6(EAAAR)$_4$ (blue), Sp1ZF6(KAAEA)$_4$ (red), and Sp1ZF6(G$_4$S)$_4$ (green) at 20 °C. (A) Free peptides in reaction buffer. (B) Free [2GC(0)] and the complexes of Sp1ZF6(EAAAR)$_4$, Sp1ZF6(KAAEA)$_4$, and Sp1ZF6(G$_4$S)$_4$ with [2GC(0)]. (C) Free [2GC(5)] and the complexes of peptide−[2GC(5)]. (D) Free [2GC(10)] and the complexes of peptide−[2GC(10)].

Sp1ZF6(EAAAR)$_4$ compared to Sp1ZF6(G$_4$S)$_4$. Generally, a minimum at 222 nm reflects the α-helix content, and then a minimum around 208 nm is a superposition of the random coil $\pi-\pi^*$ transition at 200 nm and the α-helix $\pi-\pi^*$ transition at 208 nm (*39*). The results indicated that the helical content of the linker region increases in the order of Sp1ZF6(G$_4$S)$_4$, Sp1ZF6(KAAEA)$_4$, and Sp1ZF6(EAAAR)$_4$.

Figures 2B−D show the CD spectra of the free DNA ([2GC(0)], [2GC(5)], and [2GC(10)]) and peptide−DNA complexes. The spectra of free DNA presented a pattern characteristic of the B-form of DNA (*40*), though the detailed spectra of [2GC(0)], [2GC(5)], and [2GC(10)] differ from one another because the CD spectrum of DNA is generally affected by the base composition. When the three zinc finger peptides bound to the DNA, the spectra changed even in the area of 245−320 nm, where peptides have no ellipticity, and the CD signals come from DNA (*41*). Interestingly, the CD spectra of each 6-zinc finger peptide−[2GC(0)] complex were quite similar in the area of 245−320 nm, independent of the linker sequences of the 6-zinc finger peptides (Figure 2B). The same tendency was observed in the CD spectra of the peptide−[2GC(5)] or −[2GC(10)] complexes (Figures 2C and D). It is indicated that DNA bindings of the 6-zinc finger peptides induce conformational changes in DNA and that the conformations of DNA in the complexes with three 6-zinc finger peptides are quite similar. Provided that the contributions of DNA to the CD spectra are almost equal among the complexes with the three 6-zinc finger peptides, the different spectra can be attributed to the peptides (*42*), especially to the linker structures in the 6-zinc finger peptides between the two common 3-zinc finger domains of Sp1. Table 1 shows the differences in the mean residue molecular ellipticities at 222 nm ($\Delta[\theta]_{222(\text{linker})}$), derived from the linker regions, between the DNA complexes of Sp1ZF6(EAAAR)$_4$ or Sp1ZF6(KAAEA)$_4$ and Sp1ZF6(G$_4$S)$_4$, called $\Delta[\theta]_{222(\text{EAAAR})4}$ or $\Delta[\theta]_{222(\text{KAAEA})4}$, respectively. $\Delta[\theta]_{222(\text{EAAAR})4}$ in the complexes with [2GC(10)] was comparable to that in

Table 1: Differences in Mean Residue Molecular Ellipticities at 222 nm, $\Delta[\theta]_{222(\text{linker})}$[a], Derived from the Linker Regions

	$\Delta[\theta]_{222(\text{linker})}$ (deg cm^2 dmol^{-1})	
binding site[b]	$\Delta[\theta]_{222(\text{EAAAR})4}$	$\Delta[\theta]_{222(\text{KAAEA})4}$
none	22,400	11,300
[2GC(10)]	21,500	10,200
[2GC(5)]	13,100	7,600
[2GC(0)]	12,700	8,100

[a] The differences in mean residue molecular ellipticities at 222 nm, $\Delta[\theta]_{222(\text{linker})}$, were calculated as described in the Materials and Methods section. [b] The nomenclature is described in Figure 1. The binding site None indicates the DNA-free condition.

Table 2: Comparison of Apparent Dissociation Constants (K_d)[a] for Each Peptide to Three Target DNAs, [2GC(10)], [2GC(5)], and [2GC(0)]

peptide	relative K_d[b] (K_d, nM)		
	[2GC(10)]	[2GC(5)]	[2GC(0)]
$Sp1ZF6(EAAAR)_4$	1 (1.3 ± 0.23)	16 (21 ± 2.0)	6.3 (8.2 ± 0.75)
$Sp1ZF6(KAAEA)_4$	1 (3.9 ± 0.40)	6.4 (25 ± 1.9)	5.1 (20 ± 0.63)
$Sp1ZF6(G_4S)_4$	1 (5.3 ± 0.61)	3.8 (20 ± 1.0)	2.8 (15 ± 0.19)

[a] Apparent dissociation constants were determined by titration using the gel mobility shift assay as described in the Materials and Methods section. The values are averages of three or more independent determinations with standard deviations. [b] The relative K_d values were calculated by normalizing the K_d values to each target DNA with that to [2GC(10)].

the free peptides. Interestingly, the $\Delta[\theta]_{222(EAAAR)4}$ values decreased in the complexes with [2GC(0)] and [2GC(5)] compared to the free one or in the complex with [2GC(10)]. The values of $\Delta[\theta]_{222(KAAEA)4}$ were lower than those of $\Delta[\theta]_{222(EAAAR)4}$ in accordance with the fact that the helicity of the $(KAAEA)_4$ peptide is low.

DNA Binding Affinity and Selectivity of Artificial 6-Zinc Finger Peptides. To investigate the effects of the α-helical linker on DNA binding, gel mobility shift assays were performed using $Sp1ZF6(EAAAR)_4$, $Sp1ZF6(KAAEA)_4$, and $Sp1ZF6(G_4S)_4$ for the target DNA, and [2GC(0)], [2GC(5)], and [2GC(10)] (Table 2). All of the peptides showed a single shifted band for all of the target DNA dependent on the concentrations (data not shown), suggesting the formation of a single 1:1 complex. Among the three target DNA, all of the peptides had the lowest dissociation constants (K_d) to [2GC(10)], especially the K_d value of $Sp1ZF6(EAAAR)_4$ being the lowest. Regarding the relative affinity, $Sp1ZF6(EAAAR)_4$ had a 16-fold higher affinity to [2GC(10)] than to [2GC(5)], though the relative affinities to the different target DNA were within 6.5-fold in $Sp1ZF6(G_4S)_4$ and $Sp1ZF6(KAAEA)_4$.

In order to investigate whether $Sp1ZF6(EAAAR)_4$ can distinguish [2GC(10)] containing the discontinuous two GC-box sites in the same phase, from the opposite-phased target, [2GC(5)], competition reactions were performed (Figure 3A). $Sp1ZF6(G_4S)_4$ was also examined as a control (Figure 3B). Nearly half of the peptide-bound ^{32}P-labeled [2GC(5)] remained even by addition of non-labeled [2GC(5)], whereas almost none of the peptide-bound ^{32}P-labeled [2GC(5)] remained when the non-labeled [2GC(10)] was used as the competitor (Figure 3A, lanes 2−4). Additionally, only a little decrease appeared in the peptide-bound ^{32}P-labeled [2GC(10)] by using non-labeled [2GC(5)] as the competitor, whereas the peptide-bound ^{32}P-labeled [2GC(10)] decreased by half in the case of non-labeled [2GC(10)] as the competitor (Figure 3A, lanes 6−8). Therefore, $Sp1ZF6(EAAAR)_4$ predominantly binds to [2GC(10)] in the mixture of the two target sequences, [2GC(10)] and [2GC(5)]. However, no remarkable difference was induced by $Sp1ZF6(G_4S)_4$ in the same conditions (Figure 3B).

DISCUSSION

Previous reports revealed that short peptides with $(EAAAR)_n$ ($n > 2$) sequences form α-helix structure stabilized by salt bridges and that a $(KAAEA)_3$ peptide does not (*33*, *36*). In this study, the $(EAAAR)_4$ linker was inserted between the

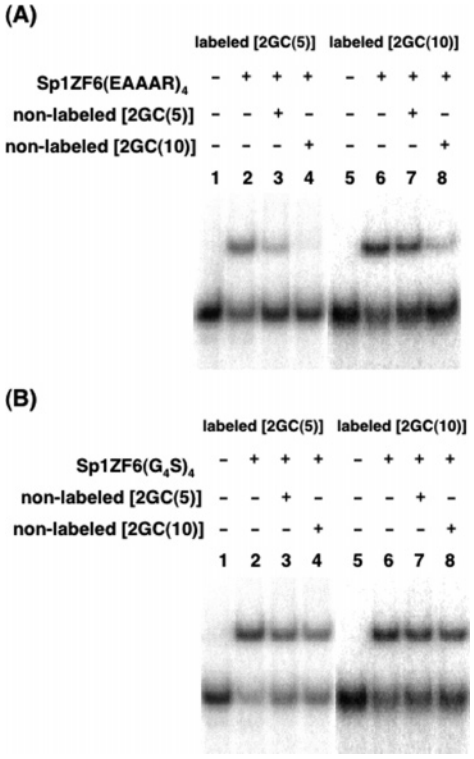

FIGURE 3: Competition experiments between [2GC(5)] and [2GC(10)] bound by $Sp1ZF6(EAAAR)_4$ (A) and $Sp1ZF6(G_4S)_4$ (B).

terminal residue of the α-helix in finger 3 and the first residue of the β-sheet in finger 4 (Figure 1A). The $(KAAEA)_4$ and $(G_4S)_4$ sequences were used as controls for the less helical and flexible linkers, respectively. The CD spectra of $Sp1ZF6(EAAAR)_4$, $Sp1ZF6(KAAEA)_4$, and $Sp1ZF6(G_4S)_4$ indicate that the α-helix content of the $(EAAAR)_4$ linker is the highest among the three linkers of 6-zinc finger peptides, as expected from the properties of linker−peptide sequences.

The gel mobility shift assays indicated that $Sp1ZF6(EAAAR)_4$ binds to [2GC(10)] with the highest affinity in the three peptides, consistent with a previous report in which the 6-zinc finger peptides with structured linkers show a higher affinity than that with a flexible linker (*23*). In addition, $Sp1ZF6(EAAAR)_4$ has the obvious preference for the [2GC(10)] sequence against [2GC(5)] in contrast with $Sp1ZF6(G_4S)_4$ as shown by the competitive experiments (Figure 3). The high affinity and preference of $Sp1ZF6(EAAAR)_4$ to [2GC(10)] should be related to the linker structures and the relative positions of the target GC-box sequences.

Considering the helical DNA structure, the two recognition sites (GC-box) are in the same phase of the DNA helix in [2GC(0)] and [2GC(10)], and in the opposite phase in [2GC(5)]. We estimated the relative positions of fingers 3 and 4 in the DNA complex by exploiting the molecular models (Figure 4). Although the complete structure of the Sp1−DNA complex has not been determined, the DNA binding mode of the Sp1 zinc finger domain is considered to be similar to that of the Zif268 zinc finger domain (*43*). Therefore, the molecular models of the 6-zinc finger−DNA complexes were constructed by inserting the gap DNA sequences between two Zif268−DNA complexes (pdb: 1AAY) (*41*). In the case of binding to [2GC(0)] with two contiguous recognition sites, the positions of the C-terminal

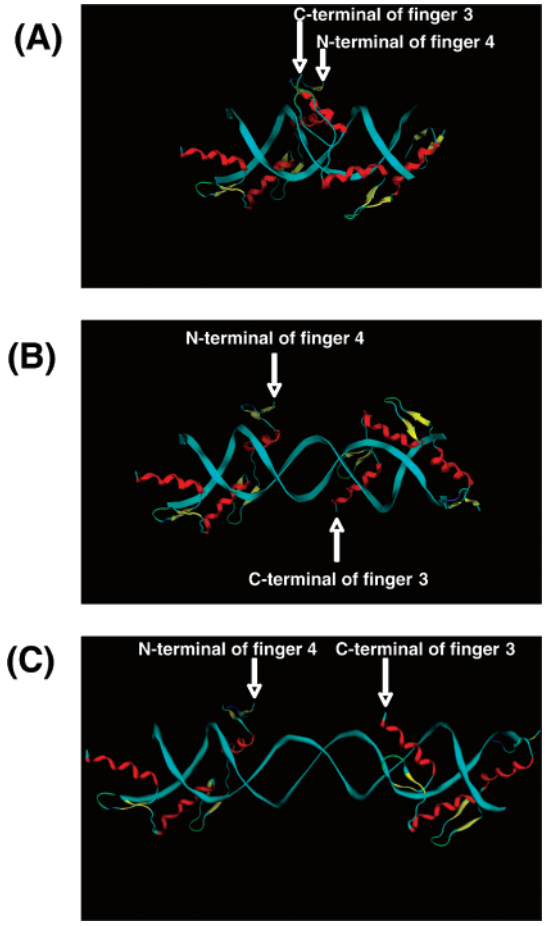

FIGURE 4: Molecular models of 6-zinc finger peptide−DNA complexes using the MOE program (Chemical Computing Group), based on the structure of the Zif268−DNA complex (*41*). The conformations of the inserted DNA gaps were set to be B-DNA. (A) The target DNA consists of continuous zinc finger binding sites (pdb 1P47) (*44*). (B) The target DNA consists of two zinc finger binding sites separated by the 5′-CTATAT-3′ nonbound gap. (C) The target DNA consists of two zinc finger binding sites separated by the 5′-CTATAATTATA-3′ nonbound gap.

of finger 3 and the N-terminal of finger 4 are expected to be close (Figure 4A). The linker is probably flipped out of the DNA major groove, and the α-helical structure of the $(EAAAR)_4$ linker should be collapsed. In the case of binding to [2GC(5)], the C-terminal of finger 3 and the N-terminal of finger 4 locate on the opposite sides of the DNA helix (Figure 4B). The linker should cross over the DNA double strand along the DNA major groove. Therefore, the linker should be extended and curved. In the complex with [2GC-(10)], the C-terminal of finger 3 and the N-terminal of finger 4 locate on the same side of the DNA helix (Figure 4C). The linker between fingers 3 and 4 should bridge a DNA minor groove along the axis of the double stranded DNA to bind to [2GC(10)]. In addition, the estimated length of the α-helical $(EAAAR)_4$ linker is close to that of one DNA helical turn. Thus, the $(EAAAR)_4$ linker is probably able to keep a high α-helix content even in the complex with [2GC-(10)]. The $\Delta[\theta]_{222(EAAAR)4}$ for [2GC(10)] suggests that the $(EAAAR)_4$ linker retains an evidently higher α-helical content in the complex with [2GC(10)] than with the flexible $(G_4S)_4$ linker. The suitable length and structural stability of the linker of Sp1ZF6$(EAAAR)_4$ for the target DNA, [2GC-(10)], could contribute to the increase in affinity rather than the other two peptides with more flexible linkers. However, $\Delta[\theta]_{222(EAAAR)4}$ decreases in the cases of [2GC(5)] and [2GC-(0)] compared to those of [2GC(10)] and the free peptide. It is indicated that the helical content of the $(EAAAR)_4$ linker decreases in the complexes with [2GC(5)] and [2GC(0)]. The same tendency was observed in Sp1ZF6$(KAAEA)_4$, though the α-helix content is lower than that in Sp1ZF6$(EAAAR)_4$. The destabilization of the linker region may result in a loss of DNA binding affinity to [2GC(0)] and [2GC(5)] compared to that to [2GC(10)]. Comparing the affinity of Sp1ZF6-$(EAAAR)_4$ to [2GC(0)] with that to [2GC(5)], a 2.5-fold higher affinity is observed. When binding to [2GC(0)], the $(EAAAR)_4$ linker is probably flipped out of the DNA major groove and significantly bent. Salt bridges between the glutamic acid and arginine residues may be formed in the bent $(EAAAR)_4$ linker and contribute to stabilization. On the contrary, when binding to [2GC(5)], the $(EAAAR)_4$ linker should be highly extended and might lose the salt bridges.

Linkers that connect individual zinc finger motifs play an important role in DNA binding. For example, the conserved TGEKP linker changes its conformation from flexible to structural in order to determine the orientation of the adjacent finger by a DNA-induced capping at the C-terminus of the α-helix in each zinc finger (*45*). Moreover, it is known that the phosphorylation of the threonine residues in the conserved linkers reduces the DNA binding affinity of the zinc finger peptide (*46*, *47*). We also explored longer linker segments between fingers 3 and 4 of the 6-zinc finger peptides in order to modify DNA binding affinity and DNA structural changes (*22*, *24*, *25*). Moore et al. introduced non-sequence-specific zinc finger motifs (a mutated finger of Zif268 and the finger 4 of TFIIIA) as a linker between fingers 3 and 4 in 6-zinc finger peptides (*23*). In general, zinc finger motifs span 3 or 4 bp. Exceptionally, finger 4 of TFIIIA bridges over the DNA minor groove without making any specific base contacts with the DNA (*48*). The introduced structured linkers are expected to selectively skip the particular spans of the DNA. However, the target selectivity to particular gaps was unexpectedly low. Thus, the structured linkers seem to be adaptable to any position of the two recognition sites. So far, a rational way to obtain a much higher selectivity has not been established. In this study, we succeeded in constructing a novel 6-zinc finger peptide with a high affinity and selectivity to [2GC-(10)] against [2GC(5)] using an α-helix as a linker between fingers 3 and 4.

Engineered zinc finger peptides that bind to the desired sequences are powerful tools to control gene expression, but all of the DNA sequences are not always recognized by the designed zinc finger peptides. In addition, specific DNA structures (e.g., bending and looping), DNA modifications, and the positions of the cis-elements are also involved in gene regulation (*31*, *32*). Therefore, it is essential for artificial zinc finger peptides to selectively recognize the relative positions of the target sites or specific DNA structures. Our results provided significant information about the linker design of novel 6-zinc finger peptides that could efficiently discriminate two separated recognition sites depending on their relative positions.

ACKNOWLEDGMENT

We thank K. Toda (Ryoka Systems Inc.) for molecular modeling using the MOE program, M. Matsumoto for

assistance with the CD measurements, and S. Negi, Y. Shiraishi, and T. Morisaki for helpful discussions and scientific advice.

REFERENCES

1. Pavletich, N. P., and Pabo, C. O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å, *Science 252*, 809−817.
2. Imanishi, M., Hori, Y., Nagaoka, M., and Sugiura, Y. (2001) Design of novel zinc finger proteins: towards artificial control of specific gene expression, *Eur. J. Pharm. Sci. 13*, 91−97.
3. Pabo, C. O., Peisach, E., and Grant, R. A. (2001) Design and selection of novel Cys2His2 zinc finger proteins, *Annu. Rev. Biochem. 70*, 313−340.
4. Beerli, R. R., and Barbas, C. F., III. (2002) Engineering polydactyl zinc-finger transcription factors, *Nat. Biotechnol. 20*, 135−141.
5. Jamieson, A. C., Miller, J. C., and Pabo, C. O. (2003) Drug discovery with engineered zinc-finger proteins, *Nat. Rev. Drug Discovery 2*, 361−367.
6. Blancafort, P., Segal, D. J., and Barbas, C. F., III. (2004) Designing transcription factor architectures for drug discovery, *Mol. Pharmacol. 66*, 1361−1371.
7. Jantz, D., Amann, B. T., Gatto, G. J. Jr., and Berg, J. M. (2004) The design of functional DNA-binding proteins based on zinc finger domains, *Chem. Rev. 104*, 789−799.
8. Beltran, A., Liu, Y., Parikh, S., Temple, B., and Blancafort, P. (2006) Interrogating genome with combinatorial artificial transcription factor libraries: asking zinc finger questions, *Assay Drug Dev. Technol. 4*, 317−331.
9. Dhanasekaran, M., Negi, S., and Sugiura, Y. (2006) Designer zinc finger proteins: tools for creating artificial DNA-binding functional proteins, *Acc. Chem. Res. 39*, 45−52.
10. Papworth, M., Kolasinska, P., and Minczuk, M. (2006) Designer zinc-finger proteins and their applications, *Gene 366*, 27−38.
11. Rebar, E. J., and Pabo, C. O. (1994) Zinc finger phage: affinity selection of fingers with new DNA-binding specificities, *Science 263*, 671−673.
12. Jamieson, A. C., Wang, H., and Kim, S.-H. (1996) A zinc finger directory for high-affinity DNA recognition, *Proc. Natl. Acad. Sci. U.S.A. 93*, 12834−12839.
13. Choo, Y., Castellanos, A., García-Hernández, B., Sánchez-García, I., and Klug, A. (1997) Promoter-specific activation of gene expression directed by bacteriophage-selected zinc fingers, *J. Mol. Biol. 273*, 525−532.
14. Greisman, H. A., and Pabo, C. O. (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites, *Science 275*, 657−661.
15. Segal, D. J., Dreier, B., Beerli, R. R., and Barbas, C. F., III. (1999) Toward controlling gene expression at will: Selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences, *Proc. Natl. Acad. Sci. U.S.A. 96*, 2758−2763.
16. Dreier, B., Beerli, R. R., Segal, D. J., Flippin, J. D., and Barbas, C. F., III. (2001) Development of zinc finger domains for recognition of the 5′-ANN-3′ family of DNA sequences and their use in the construction of artificial transcription factors, *J. Biol. Chem. 276*, 29466−29478.
17. Dreier, B., Fuller, R. P., Segal, D. J., Lund, C. V., Blancafort, P., Huber, A., Koksch, B., and Barbas, C. F., III. (2005) Development of zinc finger domains for recognition of the 5′-CNN-3′ family DNA sequences and their use in the construction of artificial transcription factors, *J. Biol. Chem. 280*, 35588−35597.
18. Liu, Q., Segal, D. J., Ghiara, J. B., and Barbas, C. F., III (1997) Design of polydactyl zinc-finger proteins for unique addressing within complex genomes, *Proc. Natl. Acad. Sci. U.S.A. 94*, 5525−5530.
19. Beerli, R. R., Segal, D. J., Dreier, B., and Barbas, C. F., III (1998) Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks, *Proc. Natl. Acad. Sci. U.S.A. 95*, 14628−14633.
20. Kamiuchi, T., Abe, E., Imanishi, M., Kaji, T., Nagaoka, M., and Sugiura, Y. (1998) Artificial nine zinc-finger peptide with 30 base pair binding sites, *Biochemistry 37*, 13827−13834.
21. Kim, J.-S., and Pabo, C. O. (1998) Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants, *Proc. Natl. Acad. Sci. U.S.A. 95*, 2812−2817.
22. Imanishi, M., Hori, Y., Nagaoka, M., and Sugiura, Y. (2000) DNA-bending finger: artificial design of 6-zinc finger peptides with polyglycine linker and induction of DNA bending, *Biochemistry 39*, 4383−4390.
23. Moore, M., Choo, Y., and Klug, A. (2001) Design of polyzinc finger peptides with structured linkers, *Proc. Natl. Acad. Sci. U.S.A. 98*, 1432−1436.
24. Imanishi, M., and Sugiura, Y. (2002) Artificial DNA-bending six-zinc finger peptides with different charged linkers: distinct kinetic properties of DNA bindings, *Biochemistry 41*, 1328−1334.
25. Imanishi, M., Yan, W., Morisaki, T., and Sugiura Y. (2005) An artificial six-zinc finger peptide with polyarginine linker: selective binding to the discontinuous DNA sequences, *Biochem. Biophys. Res. Commun. 333*, 167−173.
26. Dynan, W. S., and Tjian, R. (1983) Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II, *Cell 32*, 669−680.
27. Kadonaga, J. T., Jones, K. A., and Tjian, R. (1986) Promoter-specific activation of RNA polymerase II transcription by Sp1, *Trends Biochem. Sci. 11*, 20−23.
28. Yokono, M., Saegusa, N., Matsushita, K., and Sugiura, Y. (1998) Unique DNA binding mode of the N-terminal zinc finger of transcription factor Sp1, *Biochemistry 37*, 6824−6832.
29. Pomerantz, J. L., Wolfe, S. A., and Pabo, C. O. (1998) Structure-based design of a dimeric zinc finger protein, *Biochemistry 37*, 965−970.
30. Wolfe, S. A., Grant, R. A., and Pabo, C. O. (2003) Structure of a designed dimeric zinc finger protein bound to DNA, *Biochemistry 42*, 13401−13409.
31. Semsey, S., Virnik, K., and Adhya, S. (2005) A gamut of loops: meandering DNA, *Trends Biochem. Sci. 30*, 334−341.
32. McFall, S. M., Chugani, S. A., and Chakrabarty, A. M. (1998) Transcriptional activation of the catechol and chlorocatechol operons: variations on a theme, *Gene 223*, 257−267.
33. Merutka, G., Shalongo, W., and Stellwagen, E. (1991) A model peptide with enhanced helicity, *Biochemistry 30*, 4245−4248.
34. Paulmurugan, R., and Gambhir, S. S. (2005) Novel fusion protein approach for efficient high-throughput screening of small molecule−mediating protein-protein interactions in cells and living animals, *Cancer Res. 65*, 7413−7420.
35. Bredenbeck, J., Helbing, J., Kumita, J. R., Woolley, G. A., and Hamm, P. (2005) Alpha-helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved IR spectroscopy, *Proc. Natl. Acad. Sci. U.S.A. 102*, 2379−2384.
36. Marqusee, S., and Baldwin, R. L. (1987) Helix stabilization by $Glu^{-}\cdots Lys^{+}$ salt bridges in short peptides of *de novo* design, *Proc. Natl. Acad. Sci. U.S.A. 84*, 8898−8902.
37. Uno, Y., Matsushita, K., Nagaoka, M., and Sugiura, Y. (2001) Finger-positional change in three zinc finger protein Sp1: influence of terminal finger in DNA recognition, *Biochemistry 40*, 1787−1795.
38. Hori, Y., Suzuki, K., Okuno, Y., Nagaoka, M., Futaki, S., and Sugiura, Y. (2000) Artificial zinc finger peptide containing a novel His domain, *J. Am. Chem. Soc. 122*, 7648−7653.
39. Holzwarth, G., and Doty, P. (1965) The ultraviolet circular dichroism of polypeptides, *J. Am. Chem. Soc. 87*, 218−228.
40. Ivanov, V. I., Minchenkova, L. E., Schyolkina, A. K., and Poletayev, A. I. (1973) Different conformations of double-stranded nucleic acid in solution as revealed by circular dichroism, *Biopolymers 12*, 89−110.
41. Elrod-Erickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions, *Structure 4*, 1171−1180.
42. Talanian, R. V., McKnight, C. J., and Kim, P. S. (1990) Sequence-specific DNA binding by a short peptide dimmer, *Science 249*, 769−771.
43. Oka, S., Shiraishi, Y., Yoshida, T., Ohkubo, T., Sugiura, Y., and Kobayashi, Y. (2004) NMR structure of transcription factor Sp1 DNA binding domain, *Biochemistry 43*, 16027−16035.
44. Peisach, E., and Pabo, C. O. (2003) Constraints for zinc finger linker design as inferred from X-ray crystal structure of tandem Zif268-DNA complexes, *J. Mol. Biol. 330*, 1−7.
45. Laity, J. H., Dyson, H. J., and Wright, P. E. (2000) DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers, *J. Mol. Biol. 295*, 719−727.

46. Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B. S., and Smale, S. T. (2002) A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains, *Genes Dev. 16*, 2985−2990.
47. Jantz, D., and Berg, J. M. (2004) Reduction in DNA-binding affinity of Cys2His2 zinc finger proteins by linker phosphorylation, *Proc. Natl. Acad. Sci. U.S.A. 101*, 7589−7593.
48. Nolte, R. T., Conlin, R. M., Harrison, S. C., and Brown, R. S. (1998) Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex, *Proc. Natl. Acad. Sci. U.S.A. 95*, 2938−2943.

BI7006417