Effects of Length and Position of an Extended Linker on Sequence-Selective DNA Recognition of Zinc Finger Peptides†

Wataru Nomura and Yukio Sugiura*

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

Received August 12, 2003; Revised Manuscript Received October 18, 2003

ABSTRACT: Engineered zinc finger proteins revealed that a linker sequence connecting zinc finger units has a significant effect on the DNA binding property of the protein. The recognition for a noncontiguous DNA target beyond the current recognition code of zinc finger proteins has never been determined because of the limitation of a zinc finger framework. DNA recognition of zinc finger proteins is limited only to a contiguous subset of three base pairs. We propose the recognition for a noncontiguous DNA target by inserting amino acids into the canonical linker between zinc finger units. The sequence selectivity of the new zinc finger peptides was evaluated by gel mobility shift assays. DNase I footprinting analyses clearly showed different DNA binding of various linker-extended zinc finger peptides. The application of a SPR measurement also revealed a DNA sequence selectivity of peptides. Insertion of three amino acids is enough for recognition of a noncontiguous DNA target with sequence selectivity. An extended linker will be useful for expansion of the recognition code of zinc finger proteins and for development of a new role for linker sequences in DNA binding of zinc finger proteins.

The Cys_2His_2-type zinc finger motif is the most common DNA binding motif in eukaryotes and has tandem repeats of small units consisting of the (Tyr,Phe)-X-Cys-$X_{2,4}$-Cys-X_3-Phe-X_5-Leu-X_2-His-X_{3-5}-His-X_{2-6} consensus sequence. Each finger unit forms a compact globular structure that folds as a $\beta\beta\alpha$ module by tetrahedral coordination of a zinc ion to the invariant cysteines and histidines (*1*). DNA recognition by the Cys_2His_2-type zinc finger protein is mainly accomplished by the key amino acids at positions −1, 3, and 6 of the α-helix (*2*). To obtain the best combination of amino acids in the α-helix for the desired DNA sequence, phage display-based (*3*−*7*) or site-directed mutational (*8*, *9*) methods have been applied. Currently, the recognition code of zinc finger proteins is limited to a subset of all possible codons. To expand the recognition code of zinc finger proteins, another methodology is needed. The linker sequence between zinc finger units is highly conserved in the Cys_2His_2-type zinc finger proteins and well-known as Thr-Gly-Glu-Lys-Pro (TGEKP). The linker sequence in the free peptide is a flexible structure, but the linker sequence in the protein−DNA complex is a compact structure matched to the DNA major groove (*10*). Recently, adjustment of the linker length between two-finger units (*11*) and use of a structured linker between three-finger peptides (*12*) showed improved DNA binding. We previously demonstrated that alteration of the linker sequence has a significant effect on the DNA binding mode (*13*, *14*). However, the contribution of a linker sequence for the recognition code of zinc finger proteins has not been revealed. In the structure-based linking strategy, the modification of the linker sequence is a promising approach for sequence-selective recognition on noncontiguous DNA sequences and for expansion of the recognition code of zinc finger proteins. Here, we investigated the relationship between the length of the canonical linker and selectivity for a target DNA sequence. To achieve sequence-selective recognition on a noncontiguous DNA target, we constructed new zinc finger peptides. These peptides have extended linkers created by insertion of spacer amino acid residues. On the basis of the DNA recognition mode of zinc figner peptides, we propose that extended linker peptides can recognize the noncontiguous position in target sequences. With the extended linker mutants, we attempt to describe the effect of length and position of linker sequence on binding to contiguous and noncontiguous DNA targets. The specific goal is to show that recognition of a noncontiguous DNA target will contribute to expansion of the recognition code of zinc finger proteins.

MATERIALS AND METHODS

Chemicals. All enzymes were purchased from New England Biolabs (Beverly, MA). The synthesized oligonucleotides for construction of genes and substrate DNA were acquired from Amersham Biosciences (Cleveland, OH). The labeled compound [γ-^{32}P]ATP was supplied by DuPont. The pBS-Sp1-f1 plasmid was kindly provided by R. Tjian. All other chemicals were commercial reagent grade.

Preparation of Zinc Finger Peptides from Sp1(530−*623).* The primary structures of all the zinc finger peptides used in this study are summarized in Figure 1A. Zf123, which

† This study was supported in part by a Grant-in-Aid for COE Project "Element Science" (12CE2005) and a Grant-in-Aid for Scientific Research (13557210·14370755) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan. W.N. is a research fellow of the Japan Society for the Promotion of Science.

* To whom correspondence should be addressed: Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan. Telephone: +81-774-38-3210. Fax: +81-774-32-3038. E-mail: sugiura@scl.kyoto-u.ac.jp.

10.1021/bi035446h CCC: $25.00
Published on Web 11/26/2003

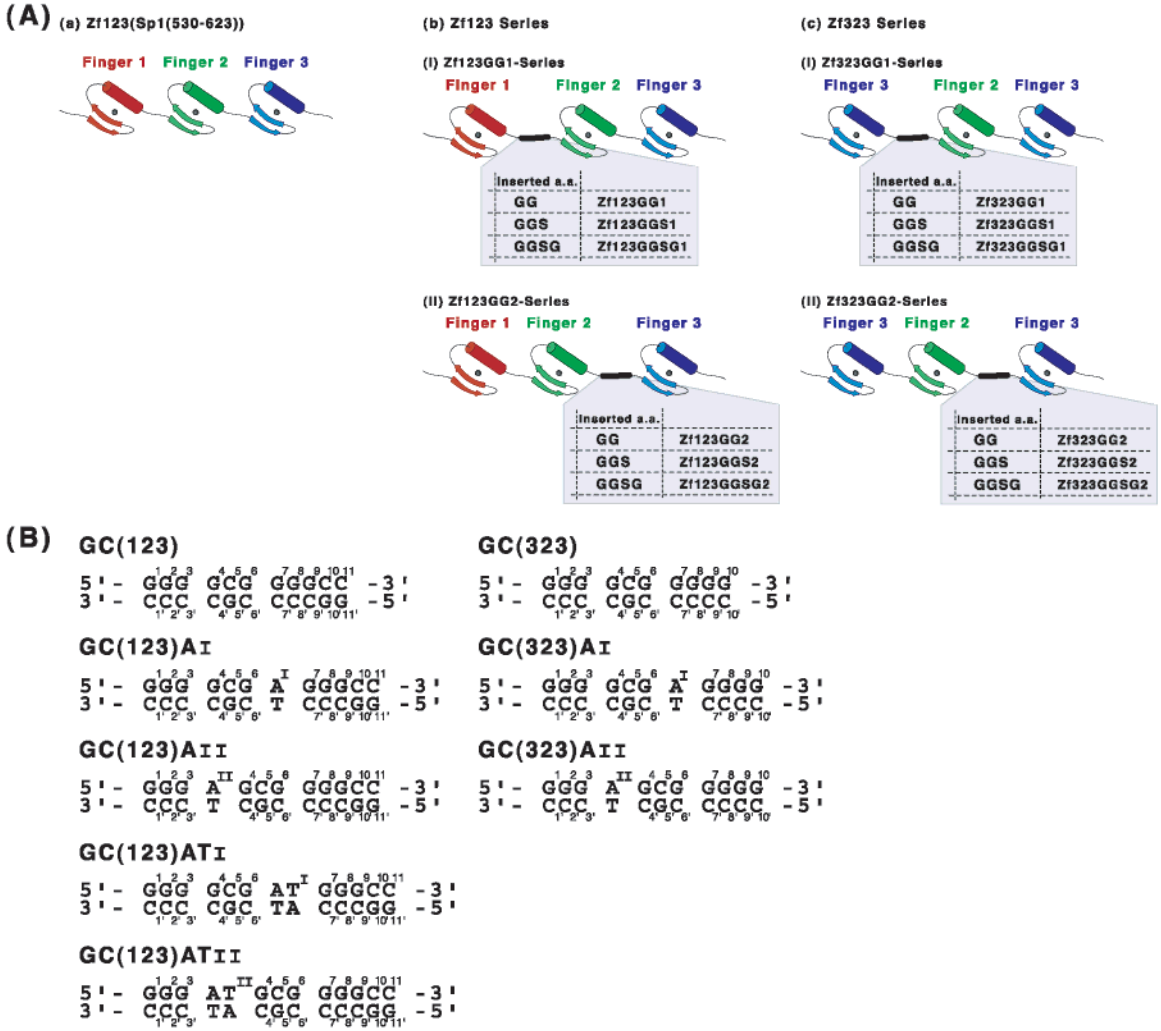

FIGURE 1: (A) Schematic drawings of Zf123 and its linker-extended mutants. (B) Sequences of GC(123) and GC(323) and the predicted sequences of target sites for the linker-extended mutants.

represents Sp1(530−623), is encoded on the pEVSp1(530−623) plasmid, as previously described (*9*, *15*). Zf323 is also identical to Sp1(zf323) prepared as previously described (*16*). All other mutant peptides were created by the standard polymerase chain reaction with the primer sets of pEVSp1(530−623) as a template for the Zf123 series and pEVSp1zf323 for the ZF323 series. Their sequences were confirmed by a GeneRapid DNA sequencer (Amersham Biosciences). These zinc finger peptides were overexpressed as a soluble form in *Escherichia coli* strain BL21(DE3)pLysS at 20 °C and purified with the following procedure at 4 °C. *E. coli* cells were resuspended and lysed in TN buffer [10 mM Tris-HCl[1] (pH 8.0), 50 mM NaCl, and 1 mM dithiothreitol]. After centrifugation, the supernatant containing the soluble form of the peptide was purified by two steps of cation exchange chromatography using a 0.05 to 2.0 M NaCl gradient (High S and Uno S, Bio-Rad, Randolph, MA). Final purification was achieved with a gel filtration technique (Superdex 75, Amersham Bioscience) using TN buffer. For substrate DNA, the *Hin*dIII−*Xba*I fragments of GC(123) (5′-GGG GCG GGG CC-3′), GC(323) (5′-GGG GCG GGG G-3′), GC(123)AI (5′-GGG GCG A GGG CC-3′), GC(123)ATI (5′-GGG GCG AT GGG CC-3′), GC(123)AII (5′-GGG A GCG GGG CC-3′), GC(123)ATII (5′-GGG AT GCG GGG CC-3′), GC(323)-AI (5′-GGG GCG A GGG G-3′), and GC(323)AII (5′-GGG A GCG GGG GC-3′) were cut out and labeled at the 5′-end with ^{32}P for the experiments as previously described (*9*).

Gel Mobility Shift Assays. Gel mobility shift assays were carried out under the previously described experimental conditions. Each reaction mixture contained 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM dithiothreitol, 10 μM $ZnCl_2$, 25 ng/μL poly(dI-dC), 0.05% Nonidet P-40, 5% glycerol, 40 mg/μL bovine serum albumin, the ^{32}P 5′-end-labeled sustrate DNA fragment (~50 pM), and 0−2000 nM zinc finger peptide. After incubation at 20 °C for 30 min, the sample was run on a 10% polyacrylamide gel with 89 mM Tris-borate buffer at 20 °C. The bands were visualized by autoradiography and quantified using ImageMaster 1D Elite software (version 3.01). The dissociation constants (K_d) of the peptide−DNA fragment complexes were evaluated by fitting the experimentally obtained values of θ_b (the fraction of labeled DNA bound to the protein) to the binding isotherm equation (eq 1) using KaleidaGraph (Abelbeck Software).

$$\theta_b = [\text{protein}]/([\text{protein}] + K_d) \quad (1)$$

[1] Abbreviations: Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; SPR, surface plasmon resonance; RU, response unit.

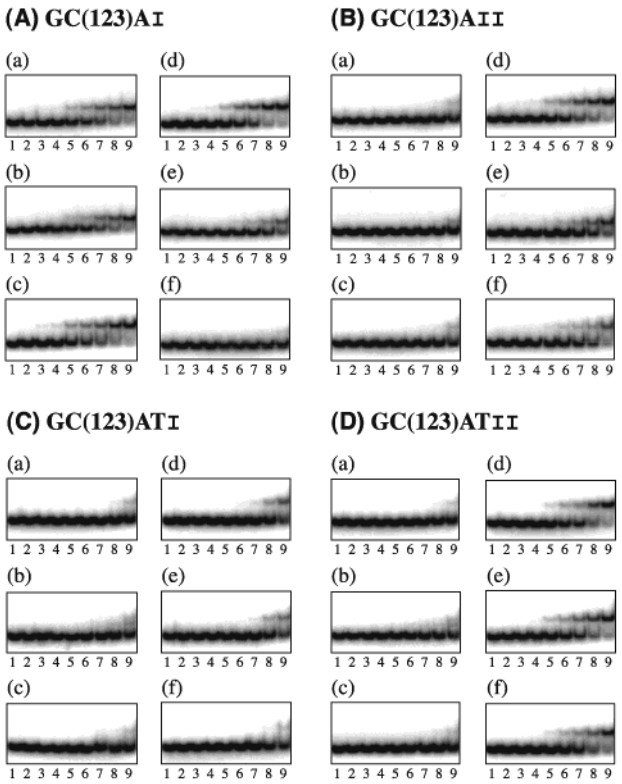

FIGURE 2: Gel mobility shift assays for the Zf123 series binding to the predicted target sequences, GC(123)AI (A), GC(123)AII (B), GC(323)ATI (C), and GC(323)ATII (D). Panels a−f depict the results of mutants Zf123GG1 (a), Zf123GGS1 (b), Zf123GGSG1 (c), Zf123GG2 (d), Zf123GGS2 (e), and Zf123GGSG2 (f). Lanes 1−9 in each panel contained 0, 100, 200, 250, 500, 750, 1000, 1500, and 2000 nM peptide, respectively.

DNase I Footprinting Analyses. DNase I footprinting experiments were performed according to the method of Brenowitz *et al*. (*17*). The binding reaction mixture contained 20 mM Tris-HCl (pH 8.0), 15 mM NaCl, 5 mM $CaCl_2$, 10 mM $MgCl_2$, 20 ng/mL sonicated calf thymus DNA, the 5′-end-labeled substrate DNA fragment (∼8 nM, 20 000 cpm), and 0−10 μM peptide. After incubation at 20 °C for 30 min, the sample was digested with DNase I (1.5 milliunit) at 20 °C for 2 min. The reaction was stopped by addition of 20 μL of DNase I stop solution (0.1 M EDTA and 0.6 M sodium acetate) and 100 μL of ethanol. After ethanol precipitation, the cleavage products were analyzed on a 15% polyacrylamide/7 M urea sequencing gel. The bands were visualized by autoradiography.

SPR Measurement. The operational principle of the BIACORE biosensor has already been described (*18*). The protein−DNA interaction was studied using a BIACORE X instrument (BIACORE AB, Uppsala, Sweden) operated at 25 °C. The 5′-biotinylated oligonucleotide (5′-biotin-TGGATCTGGGGCGAGGGGTAATTCG-3′) was annealed with the complementary strand in flow buffer. The duplex DNA solution was injected over a streptavidin-coated sensor chip (SA5, BIACORE AB) until a suitable level (200−300 RU) was achieved. Flow cell 1 was left unmodified as the control. Tris-HCl buffer [10 mM Tris-HCl (pH 7.7), 250 mM NaCl, 20 mM $MgCl_2$, 0.1 mM $ZnCl_2$, and 0.005% Tween 20] was used as both the flow buffer and the sample preparation buffer. The peptide concentrations were 125, 375, and 750 nM. The association was followed for 5 min and the dissociation for 10 min at a flow rate of 20 μL/min. The bound protein was eluted from DNA by several repeats of a short pulse (5 μL) of the regeneration solution (0.02% SDS in 50 mM NaOH). Analysis of the data was performed using the evaluation software supplied with the instrument (BIAevaluation, version 3.0).

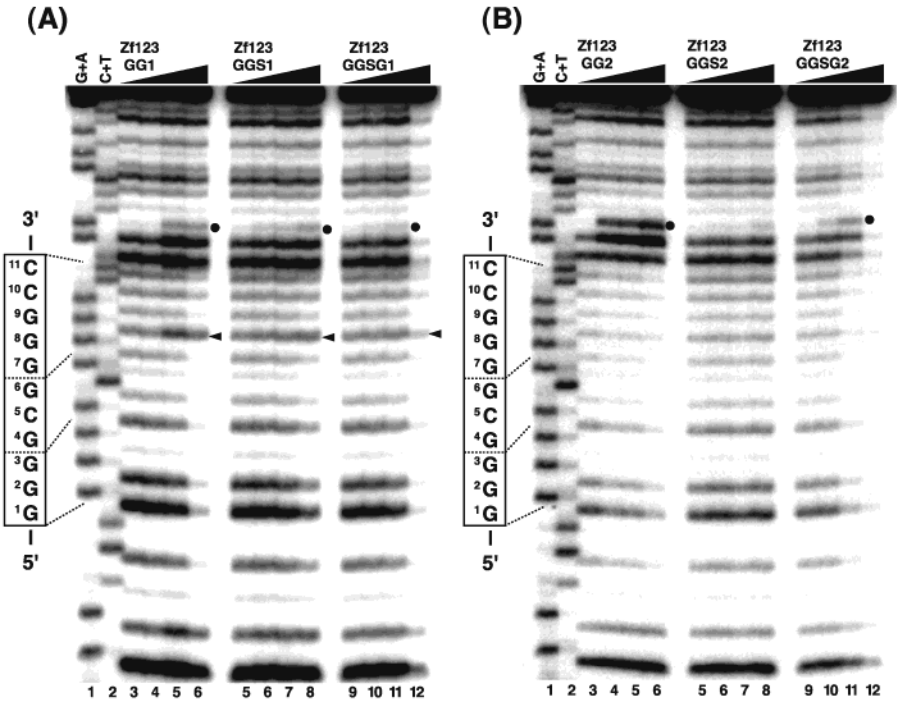

FIGURE 3: DNase I footprinting analyses of the Zf123 series for binding to GC(123). Panels A and B show the results of the Zf123GG1 and Zf123GG2 series, respectively: lane 1, G+A (Maxam−Gilbert reaction products); lane 2, C+T (Maxam−Gilbert reaction products); lanes 3, 5, and 9, no peptide; lanes 4, 6, and 10, 1.0 μM peptide; lanes 5, 7, and 11, 2.5 μM peptide; and lanes 6, 8, and 12, 5.0 μM peptide. The dots denote the hypersensitive cleavage sites, and the arrowheads denote the unprotected cleavage sites.

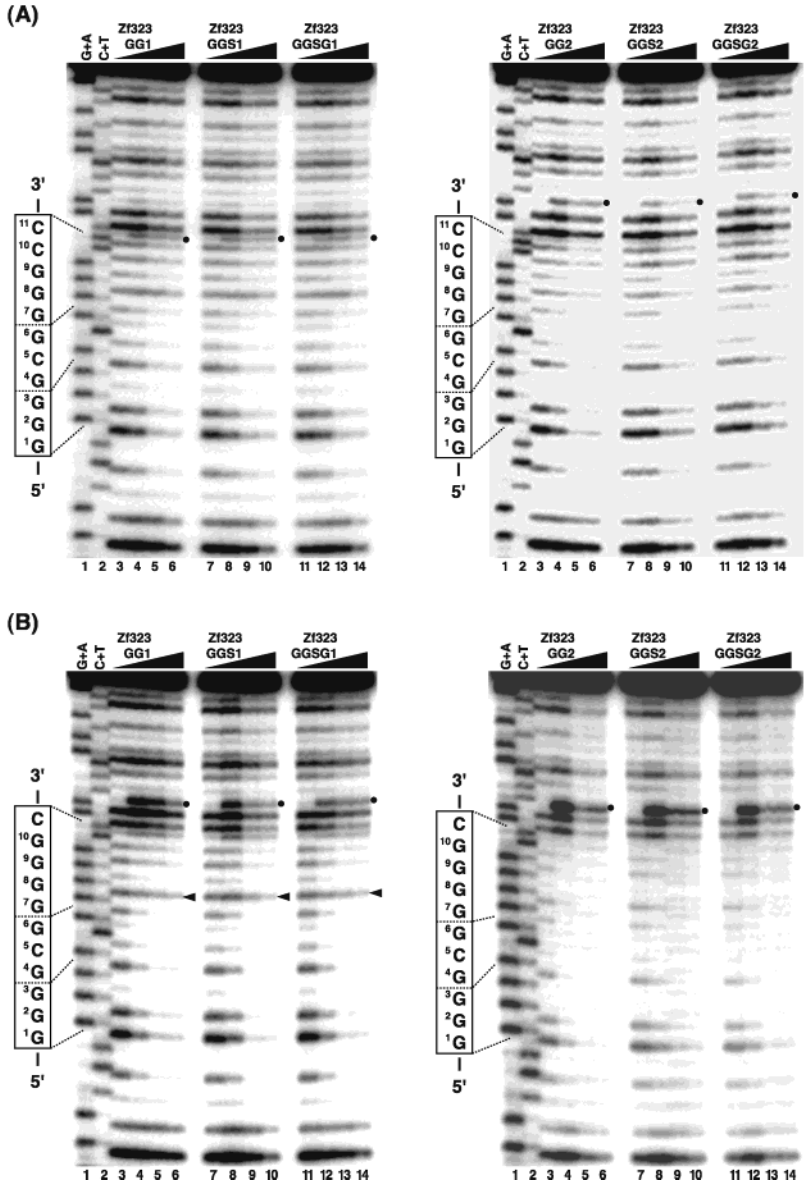

FIGURE 4: DNase I footprinting analyses of the Zf323 series for binding to GC(123) (A) and GC(323) (B): lane 1, G+A (Maxam−Gilbert reaction product); lane 2, C+T (Maxam−Gilbert reaction product); lanes 3, 7, and 11, no peptide; lanes 4, 8, and 12, 2.5 μM peptide; lanes 5, 9, and 13, 5.0 μM peptide; and lanes 6, 10, and 14, 10 μM peptide. The dots denote the hypersensitive cleavage sites, and the arrowheads denote the unprotected cleavage sites.

RESULTS

Designs of the Linker-Extended Zinc Finger Peptides, Zf123 Series, Zf323 Series, and Their Predicted Binding Sequences. To study the effect of extension of a linker sequence between zinc finger units on DNA binding, 12 three-zinc finger mutants were prepared. On the basis of the Gly-Gly-Ser linker that is generally used as a spacer sequence between protein domains, the three kinds of amino acid sequences were used for extension of the linker sequence. These linker sequences exclude the interaction between amino acids. Then, only the effect of the linker length on DNA binding can be estimated. Zf123GG1 series Zf123GG1/GGS1/GGSG1 contain Gly-Gly/Gly-Gly-Ser/Gly-Gly-Ser-Gly extended linker sequences between fingers 1 and 2 (F1−F2 linker) and Zf123GG2 series Zf123GG2/GGS2/GGSG2 between fingers 2 and 3 (F2−F3 linker). The same mutants for Zf323 are renamed Zf323GG1/GGS1/GGSG1 and Zf323GG2/GGS2/GGSG2 (Figure 1A). Amino acids for linker extension are inserted between E3 (the glutamic acid at the third position of the linker) and the fourth residue in the canonical linker. The position of linker insertion was chosen to prevent disruption of the hydrogen bonding between the backbone amide of E3 and side chain O^{γ} of T1 (the threonine at the first position of the linker) which generates additional stabilization on protein−DNA interaction (*19*). The contiguous and noncontiguous DNA targets for the linker-extended mutants were also constructed (Figure 1B). In this study, an A·T base pair was used as a noncontiguous position because it is out of the recognition code of the Sp1 zinc finger protein. GC(123)AI contains an A·T base pair in the 5′-G^6G^7-3′ step of GC(123). For GC(123)AII, in the same way, an A·T base pair was inserted in the 5′-G^3G^4-3′ step. GC(323)AI and GC(323)AII were also constructed in the same way on the basis of GC(323). As for GC(123)ATI and GC(123)ATII, they were constructed in an effort to investigate whether linker-extended mutants

Table 1: Dissociation Constants (K_d) for the Zf123 Series Binding to GC(123)

peptide	K_d (nM)[a]	peptide	K_d (nM)[a]
Zf123GG1	133 ± 4.2	Zf123GGS2	185 ± 1.5
Zf123GGS1	247 ± 56	Zf123GGSG2	88.9 ± 15
Zf123GGSG1	245 ± 13	Zf123	14.0 ± 1.6[b]
Zf123GG2	22.9 ± 2.1		

[a] Apparent dissociation constants are determined using gel mobility shift assays as described in Materials and Methods. Values are averages of three or more independent determinations with standard deviations. [b] This value is derived from the published data in ref *8*.

retain a binding affinity for two-base pair incorporation of A·T base pairs. Previously, we reported that the two-zinc finger variants, finger 1−2 and finger 2−3 of the Sp1 zinc finger peptide, retain DNA binding activity (*9*). To prove whether the one-finger unit separated by the extended linker sequence contributes to DNA binding, the N-terminal finger substituted mutants (Zf323 series) were constructed. These constructions were based on Sp1(zf323), in which the N-terminal finger unit is substituted from finger 1 to finger 3.

DNA Binding Affinity of the Zf123 Series for a Contiguous DNA Target. The DNA binding affinities of the Zf123 series are shown in Table 1. The different effect of the linker length was observed between the Zf123GG1 series and the Zf123GG2 series. Zf123GG1 and Zf123GG2 exhibited the highest binding affinity in the series. In the Zf123GG1 series, Zf123GGS1 and Zf123GGSG1 had similar DNA binding affinities. However, in the Zf123GG2 series, Zf323GGSG2 exhibited a 2-fold higher affinity than Zf323GGS2. To compare the binding of the same linker length mutants to that of GC(123), the Zf123GG2 series evidently indicated a higher affinity than the Zf123GG1 series.

Sequence-Selective Binding for a Noncontiguous DNA Target of the Zf123 Series. The gel mobility shift assays demonstrated a clear binding selectivity for noncontiguous DNA targets (Figure 2). The K_d values for the binding of Zf123GG1, Zf123GGS1, and Zf123GGSG1 to GC(123)AI were 1.53, 1.34, and 1.21 μM, respectively. Zf123GG2 bound to GC(123)AI and GC(123)AII with the same affinity, estimated to be 1.30 μM. The K_d values for the binding of Zf123GG2, Zf123GGS2, and Zf123GGSG2 to GC(123)ATII were 1.65, 1.70, and 2.30 μM, respectively. In the binding to GC(123)ATI, all mutants exhibited poor binding.

Detailed Binding for Contiguous DNA Targets and the Recognition Mode of the Zf123 Series. The DNA binding mode of the linker-extended mutants was determined by DNase I footprinting analyses. The Zf123GG1 series exhibited a slight footprinting upon binding to GC(123) (Figure 3A). Hypersensitive cleavage at the 3′-external sequence of GC(123) was also observed. The cleavage band that represents the uncovered DNA part in the peptide binding appeared at the 5′-G^{6}G^7-3′ step of GC(123). On the other hand, the Zf123GG2 series displayed different footprinting patterns by length of the extended linker (Figure 3B). Zf123GG2 depicted a strong footprinting and hypersensitive cleavage at the 3′-external sequence of GC(123). A similar footprinting pattern was also observed in the binding of Zf123GGSG2. However, the poor footprinting was observed at a low peptide concentration in the binding of Zf323GGSG2. The footprinting pattern of Zf123GGS2 was different from those of the former two peptides. The hypersensitive DNA cleavage and the footprinting at low peptide concentrations were not detected.

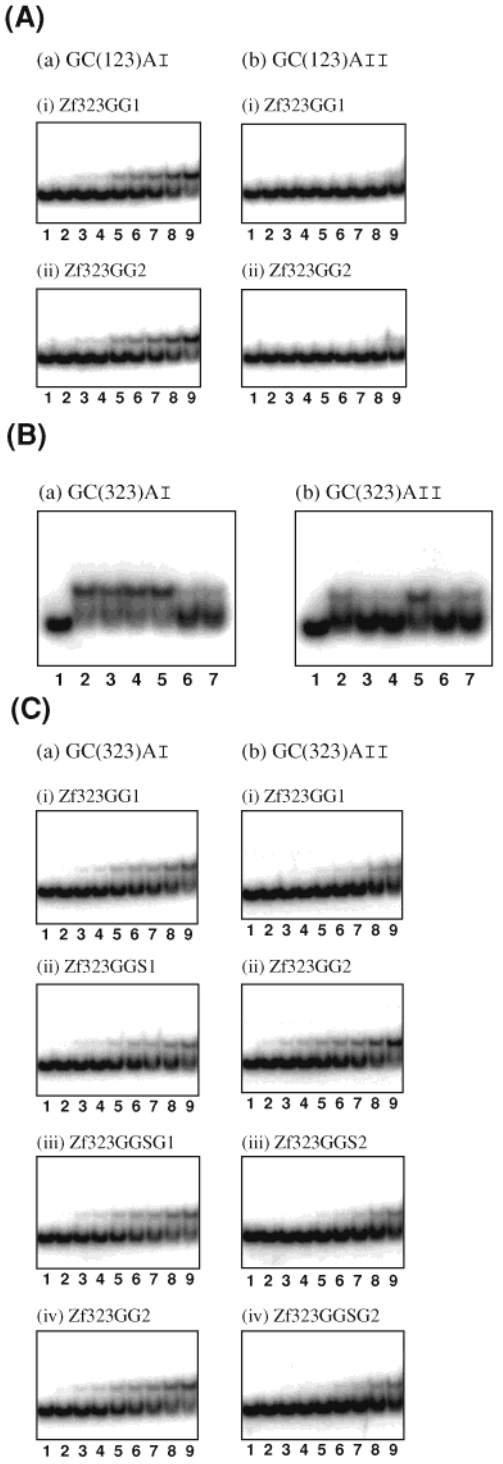

FIGURE 5: Gel mobility shift assays for the Zf323 series. (A) Binding to GC(123)AI (a) and GC(123)AII (b). Panels i and ii show the results for Zf323GG1 (i) and Zf323GG2 (ii). Lanes 1−9 in each panel contained 0, 100, 200, 250, 500, 750, 1000, 1500, and 2000 nM peptide, respectively. (B) Binding to GC(323)AI (a) and GC(323)AII (b) at 2000 nM peptide. Lanes 1−7 in each panel show the results for no peptide (lane 1), Zf323GG1 (lane 2), Zf323GGS1 (lane 3), Zf323GGSG1 (lane 4), Zf323GG2 (lane 5), Zf323GGS2 (lane 6), and Zf323GGSG2 (lane 7). (C) Binding to the high-affinity binding sequence of the Zf323 series at various peptide concentrations. The peptides are indicated. Lanes 1−9 in each panel contained 0, 100, 200, 250, 500, 750, 1000, 1500, and 2000 nM peptide, respectively.

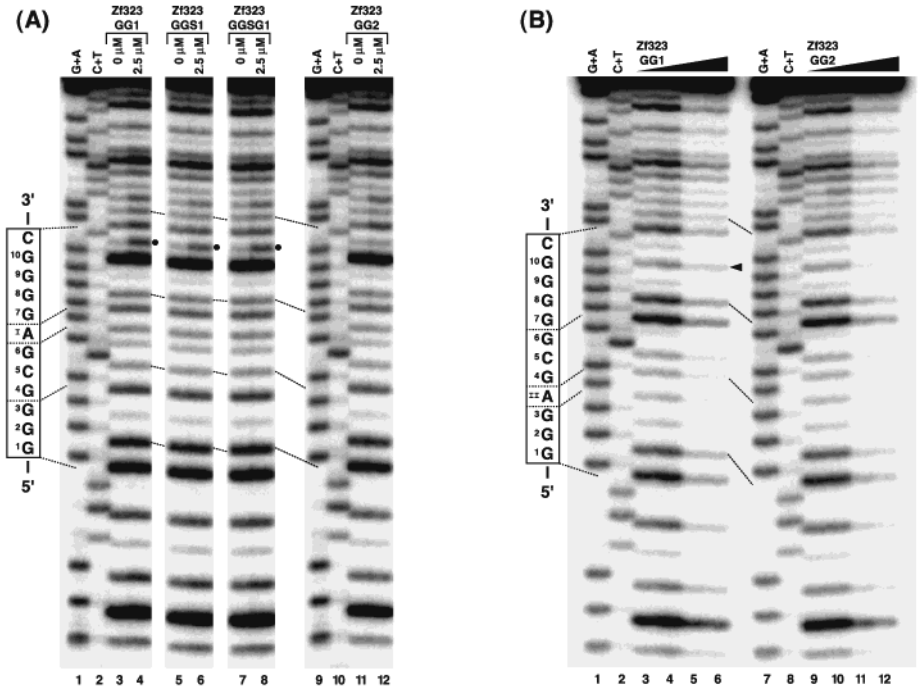

FIGURE 6: DNase I footprinting analyses of the Zf323 series for binding to GC(323)AI (A) and GC(323)AII (B). (A) Lane 1, G+A (Maxam−Gilbert reaction product); and lane 2, C+T (Maxam−Gilbert reaction product). The peptides and their concentrations are indicated. (B) Lanes 1 and 7, G+A (Maxam−Gilbert reaction product); lanes 2 and 8, C+T (Maxam−Gilbert reaction product); lanes 3 and 9, no peptide; lanes 4 and 10, 2.5 μM peptide; lanes 5 and 11, 5.0 μM peptide; and lanes 6 and 12, 10 μM peptide. The dots denote the hypersensitive cleavage sites, and the arrowheads denote the unprotected cleavage sites.

Table 2: Dissociation Constants (K_d) for the Zf323 Series Binding to GC(123) and GC(323)

	K_d (nM)[a]	
peptide	GC(123)	GC(323)
Zf323GG1	187 ± 13	76.4 ± 4.7
Zf323GGS1	252 ± 15	118 ± 12
Zf323GGSG1	285 ± 65	137 ± 37
Zf323GG2	89.2 ± 4.4	26.3 ± 8.4
Zf323GGS2	423 ± 9.0	63.9 ± 3.3
Zf323GGSG2	464 ± 98	62.7 ± 9.4

[a] Apparent dissociation constants are determined using gel mobility shift assays as described in Materials and Methods. Values are averages of three or more independent determinations with standard deviations.

Effect of N-Terminal Finger Substitution on Peptide Binding to Contiguous DNA Targets. The dissociation constants (K_d) of the Zf323 series for binding to GC(123) and GC(323) were estimated by gel mobility shift assays (Table 2). The Zf323 series obviously exhibited a higher affinity for GC(323) than for GC(123). DNase I footprinting analyses for the Zf323 series were performed to reveal the mode of binding to contiguous DNA targets. In the binding to GC(123), the Zf323GG1 and Zf323GG2 series had different hypersensitive cleavages (Figure 4A). In the binding of the Zf323GG1 series to GC(323), a similar footprinting pattern was observed in all three mutants. Interestingly, the cleavage band that represents the uncovered DNA part caused by the peptide binding was detected in the 5′-G^{6}G^7-3′ step of GC(323). In the binding of the Zf323GG2 series to GC(323), such a cleavage band was not observed (Figure 4B).

Binding to Noncontiguous DNA Targets of the N-Terminal Finger Substituted Mutants. In the binding to GC(123)AI and GC(123)AII, only Zf323GG1 and Zf323GG2 were bound to GC(123)AI (Figure 5A) and other peptide binding was not observed (data not shown). On the other hand, a clear sequence selectivity was noted in the binding to GC(323)AI and GC(323)AII (Figure 5B). The detailed binding results are depicted in Figure 5C. The K_d values for the binding of Zf323GG1, Zf323GGS1, Zf323GGSG1, and Zf323GG2 to GC(323)AI were 2.40, 3.62, 3.64, and 3.13 μM, respectively. In the binding to GC(323)AII, only Zf323GG2 bound reasonably well at the K_d value of 2.10 μM. DNase I footprinting analyses revealed the detailed mode of binding of some Zf323 mutants to GC(323)AI and GC(323)AII. The Zf323GG1 series produced the hypersensitive cleavage at the 3′-external sequence of GC(323)AI, but Zf323GG2 did not (Figure 6A). In the binding to GC(323)AII, only Zf323GG1 exhibited an interesting cleavage band in the 5′-G^{8}G^9-3′ step of GC(323)AII (Figure 6B).

Detection of Sequence-Selective DNA Binding Using SPR Measurement. SPR measurement with the BIACORE instrument has been utilized for the assessment of the protein−DNA interaction, including several zinc finger complexes with specific DNA (*20*−*22*). The binding of the Zf323 series to GC(323)AI was assessed with this measurement. The RU value represents the strength of the protein−DNA interaction. In this experiment, the final RU value in the association phase (0−5 min) was used as the value for the strength of the protein−DNA interaction. Zf323GG1 displayed a higher final RU than Zf323GG2 at concentrations of 375 and 750 nM (Figure 7A). The final RU value of Zf323GGS1 and Zf323GGSG1 was higher than that of Zf323GGS2 and Zf323GGSG2 at every concentration (Figure 7B,C). In particular, the final RU value in the association phase of Zf323GGSG1 was 2-fold higher than that of Zf323GGSG2 at 750 nM peptide. The sensorgrams of the Zf323GG1 series were different according to the numbers of inserted amino

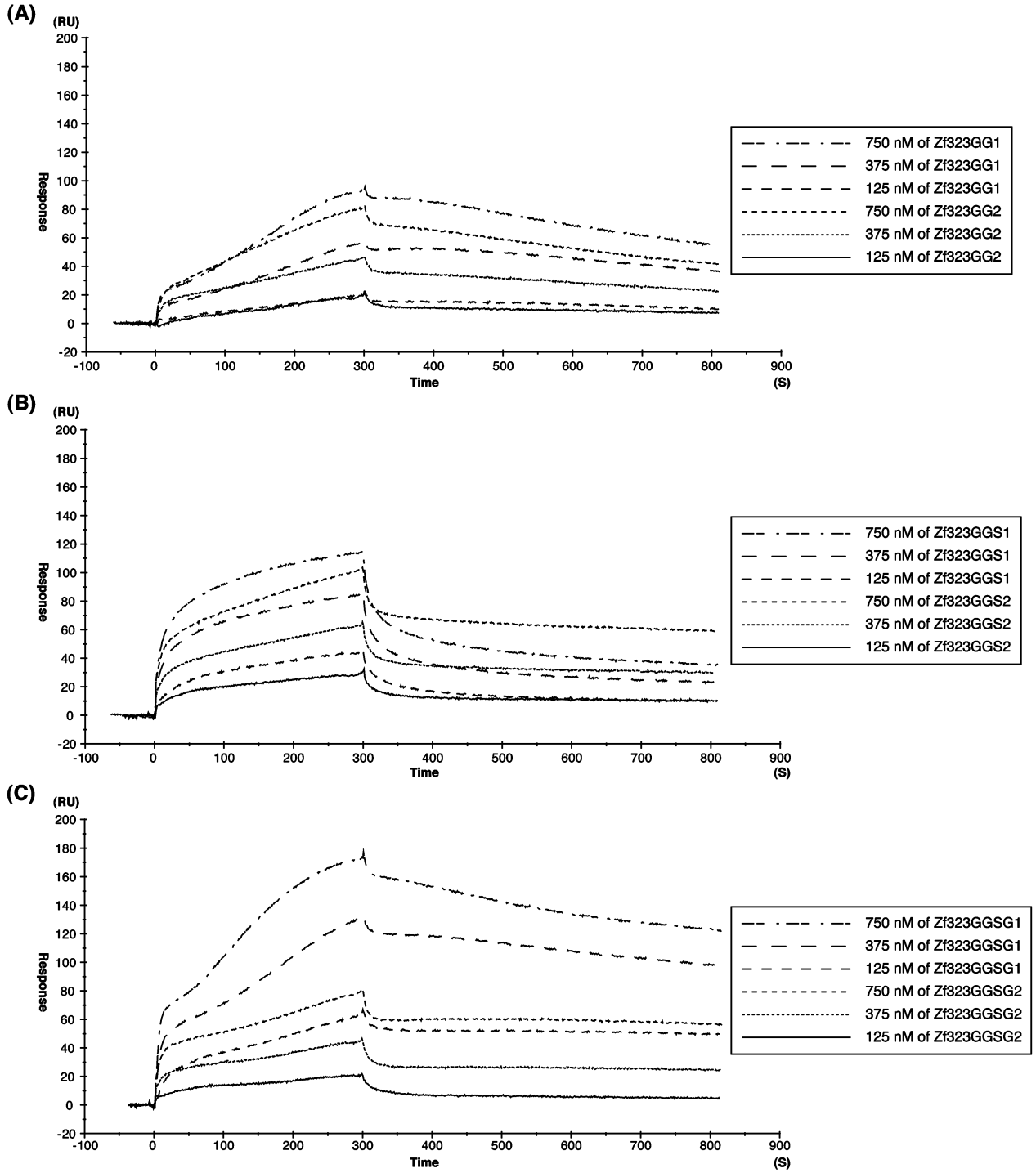

FIGURE 7: Typical response features (sensorgrams) of the binding to GC(323)AI immobilized on an SPR sensor. The binding of Zf323GG1 and Zf323GG2 (A), Zf323GGS1 and Zf323GGS2 (B), and Zf323GGSG1 and Zf323GGSG2 (C) are shown. The peptide concentrations are 125, 375, and 750 nM.

acids. The sensorgram of Zf323GGS1 was typical, like that observed for Sp1(530−623) (Figure 7B) (*23*). On the other hand, Zf323GG1 and Zf323GGSG1 exhibited sigmoid sensorgrams (Figure 7A,C). This difference appeared in odd or even numbers of inserted amino acids.

DISCUSSION

Effects of Linker Extension Caused by Insertion of Spacer Sequences. The binding affinities of the Zf123 series for binding to GC(123) were different according to the length of the linker, although those of Zf123GGS1 and Zf123GGSG1 were almost same. This result shows that the effect of linker extension is different depending on the position of linker insertion. Zinc finger proteins have finger−finger interaction. This interaction has been found in the NMR study of the DNA complex of the TFIIIA zinc finger (*24*). The difference in the effect of the position of linker insertion could be caused by the difference in the finger−finger interaction between fingers 1 and 2 and between fingers 2 and 3.

In the binding to noncontiguous DNA targets, the Zf123GG1 series prefers GC(123)AI and some mutants of the Zf123GG2 series favor GC(123)AII and GC(123)ATII.

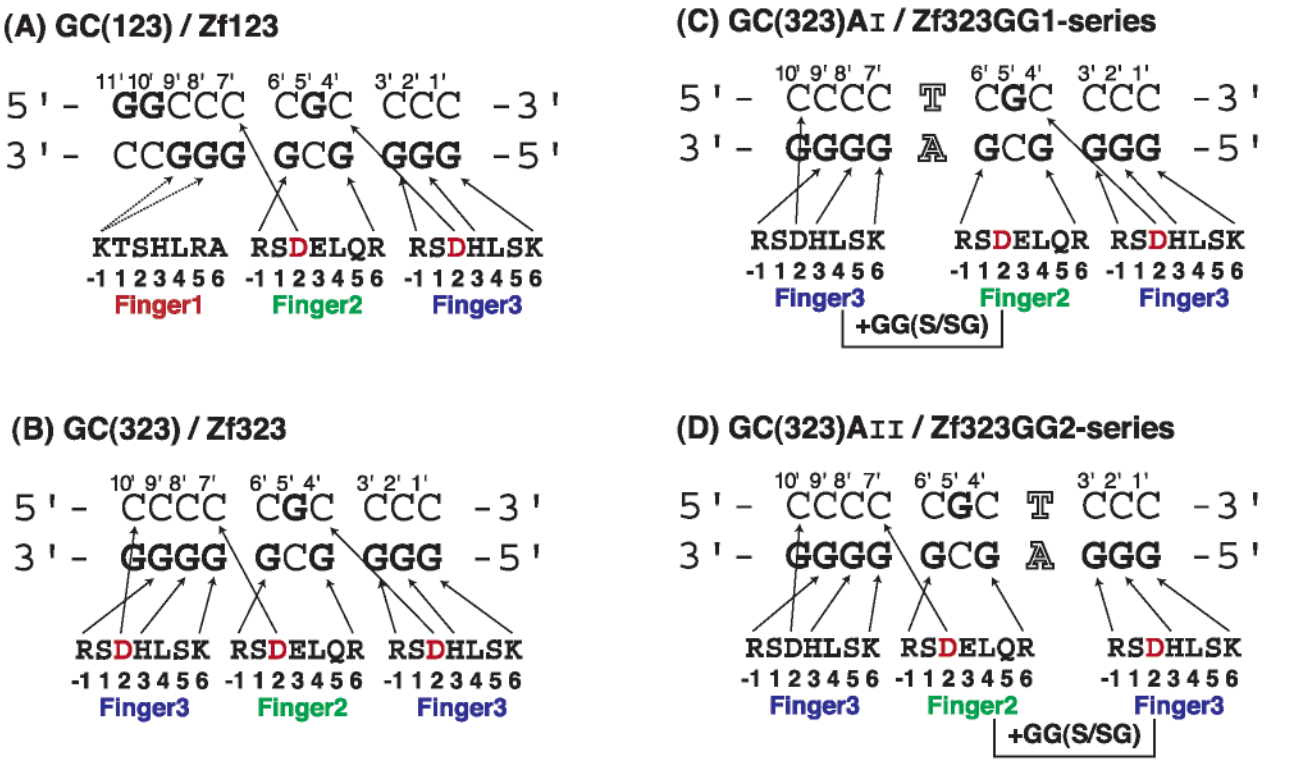

FIGURE 8: Possible base recognition of the zinc finger peptides. The interactions between the amino acids in the α-helix and DNA bases are shown: Zf123−GC(123) (A), Zf323−GC(323) (B), Zf323GG1 series−GC(323)AI (C), and Zf323GG2 series−GC(323)AII (D).

However, Zf123GG2 did not show a clear distinction in the binding to the noncontiguous DNA targets. This result indicates that the insertion of two amino acids into the F2−F3 linker still produces behavior that is the same as that of Zf123 on DNA binding. Zf123 showed poor sequence selectivity in the binding to GC(123)AI and GC(123)AII (data not shown). The Zf123 series revealed that insertion of three amino acids is enough to distinguish the sequence variety of noncontiguous DNA targets. The insertion of other amino acids would result in a different preference for the numbers of amino acids for insertion and display additional functions on DNA binding.

The linker-extended mutants exhibited relatively lower binding affinities than the original zinc fingers, Zf123 and Zf323, in the binding to noncontiguous DNA targets. This result is supposed to maintain the relationship with recognition for the cross-strand that specifically recognizes a subsite for an adjacent finger at the opposite strand. The interaction of Asp2 in the α-helix with the cross-strand bases has been assessed by phage display selections (*25*, *26*). For the Sp1 zinc finger, fingers 2 and 3 have Asp at position 2 of the helix and they contact C4′ and C7′ of GC(123) (Figure 8A); these interactions are also expected for Zf323 (Figure 8B) (*8*). During recognition for noncontiguous DNA targets, this interaction is probably disrupted by insertion of A·T base pairs (Figure 8C,D). In the Zif268 zinc finger, the loss of such an interaction by the substitution of Asp2 with Ala resulted in the decrease of the 4-fold binding affinity (*26*).

Evidence for Three-Finger Binding Revealed by N-Terminal Finger Substitution. The N-terminal finger substitution clearly provides the synergism of three finger units for binding to noncontiguous DNA targets. Zf323GG1 and Zf323GG2 bind to GC(123)AI despite the different subsite for the N-terminal finger unit. This fact indicates that the contact of two finger units maintains the DNA binding affinity. Additionally, these two peptides showed poor DNA sequence selectivity in the binding to GC(323)AI and GC(323)AII. Zf323GGS1 and Zf323GGSG1 exhibited distinct sequence selectivity for GC(323)AI. Compared with the binding of the Zf323GG1 series to GC(323)AI, the Zf323GG2 series displayed relatively poor binding to GC(323)AII except for Zf323GG2. This effect is probably caused by the loss of the cross-strand contact of the one-finger unit of the Zf323GG2 series. For the Zf323GG1 series, the separated finger unit is finger 3. In the binding to GC(323)AI, the 5′-GGGG-3′ subsite is located at the 3′-end of GC(323)AI for recognition by finger 3 (Figure 8C). On the contrary, for the binding of the Zf323GG2 series to GC(123)AII, the overlapping base is thymine, which is not recognized by Asp2 in the α-helix (Figure 8D). These results suggest that binding by the synergism of three finger units and insertion of three amino acids are needed for the recognition of noncontiguous DNA targets with sequence selectivity. To improve sequence-selective DNA recognition, the detailed synergistic binding by the adjacent fingers must be elucidated.

DNA Binding Mode of Linker-Extended Zinc Finger Peptides. Some mutants of the Zf123GG1 and Zf323GG1 series displayed an interesting cleavage band that indicates the uncovered DNA part in peptide binding. Of particular interest is the fact that this cleavage band is observed in the binding to specific DNA targets. These targets are contiguous at the position where the extended linker is situated on DNA binding. In contrast to it, in the binding to noncontiguous targets at this position, this cleavage band was not observed. Then, this cleavage band seems to maintain the relationship to the DNA binding mode of extended linker mutants.

We investigated the contribution to recognition site selectivity of linker length between zinc fingers. In this work, two- to four-amino acid residue linkers were inserted between the zinc fingers of three-finger peptides, and DNA binding activities were assayed on contiguous and noncontiguous substrates. Some of these constructs recognized the noncontiguous DNA. The results show that varying the length of the linkers between zinc fingers can be used to expand the recognition code of zinc finger proteins, which is currently limited to a subset of all possible codons.

REFERENCES

1. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., and Wright, P. E. (1989) *Science 245*, 635−637.
2. Pavletich, N. P., and Pabo, C. O. (1991) *Science 252*, 809−817.
3. Reber, E. J., and Pabo, C. O. (1994) *Science 263*, 671−673.
4. Choo, Y., and Klug, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11163−11167.
5. Choo, Y., and Klug, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11168−11172.
6. Elrod-Erickson, M., and Pabo, C. O. (1999) *J. Biol. Chem. 274*, 19281.
7. Segal, D. J., and Barbas, C. F., III (2000) *Curr. Opin. Chem. Biol. 4*, 34−39.
8. Nagaoka, M., Shiraishi, Y., Uno, Y., Nomura, W., and Sugiura, Y. (2002) *Biochemistry 41*, 8819−8825.
9. Yokono, M., Saegusa, N., Matsushita, K., and Sugiura, Y. (1998) *Biochemistry 37*, 6824−6832.
10. Laity, J. H., Dyson, H. J., and Wright, P. E. (2000) *J. Mol. Biol. 295*, 719−727.
11. Moore, M., Choo, Y., and Klug, A. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 1437−1441.
12. Moore, M., Choo, Y., and Klug, A. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 1432−1436.
13. Nagaoka, M., Nomura, W., Shiraishi, Y., and Sugiura, Y. (2001) *Biochem. Biophys. Res. Commun. 282*, 1001−1007.
14. Nomura, W., Nagaoka, M., Shiraishi, Y., and Sugiura, Y. (2003) *Biochem. Biophys. Res. Commun. 300*, 87−92.
15. Nagaoka, M., and Sugiura, Y. (1996) *Biochemistry 35*, 8716−8768.
16. Uno, Y., Matsushita, K., Nagaoka, M., and Sugiura, Y. (2001) *Biochemistry 40*, 1787−1795.
17. Brenowitz, M., Senear, D. F., Shea, M. A., and Acker, G. K. (1986) *Methods Enzymol. 130*, 132−181.
18. Nice, E. C., and Catimel, B. (1999) *BioEssays 21*, 339−352.
19. Laity, J. H., Chung, J., Dyson, H. J., and Wright, P. E. (2000) *Biochemistry 39*, 5341−5348.
20. Imanishi, M., and Sugiura, Y. (2002) *Biochemistry 41*, 1328−1334.
21. Young, E., Kacherovsky, N., and Cheng, C. (2000) *Biochemistry 39*, 567−574.
22. Schaufler, L. E., and Klevit, R. E. (2003) *J. Mol. Biol. 329*, 931−939.
23. Kamiuchi, T., Abe, E., Imanishi, M., Kaji, T., Nagaoka, M., and Sugiura, Y. (1998) *Biochemistry 37*, 13827−13834.
24. Foster, M. P., Wuttke, D. S., Radhakrishnan, I., Case, D. A., Gottesfeld, J. M., and Wright, P. E. (1997) *Nat. Struct. Biol. 4*, 605−608.
25. Isalan, M., Choo, Y., and Klug, A. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5617−5621.
26. Isalan, M., Klug, A., and Choo, Y. (1998) *Biochemistry 37*, 12026−12033.

BI035446H