Factors Governing the Sequence-Selective DNA Binding of Geometrically Constrained Peptide Dimers

Takashi Morii,* Yoshitaka Saimei,† Masako Okagami,† Keisuke Makino,† and Yukio Sugiura*

Contribution from the Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan, and Department of Polymer Science & Engineering, Kyoto Institute of Technology, Sakyo-ku, Kyoto 606, Japan

Received September 18, 1996[⊗]

Abstract: Peptide dimers of the basic leucine zipper protein with non-native monomer arrangements were synthesized by using C_2 chiral templates as a synthetic dimerization module. The amino acid sequence of the peptide is derived from the DNA contact region of the basic leucine zipper protein GCN4. These peptide dimers are designed to possess different geometrical constraints from that of native GCN4 with respect to the orientation of two DNA-contacting peptides. Peptide dimers constrained at the 6th position from N-terminus recognized novel palindromic DNA sequences in which the polarity of each half-site of the parent GCN4 binding sequence is reversed. This is in contrast with dimers that are constrained at the N-terminus which failed to recognize the reversed DNA sequences. Sequence-specific recognition of these palindromic DNA sequences was confirmed by DNase I footprinting. Circular dichroism spectral analyses revealed that dimers constrained at the 6th position bind in the helical conformation to the reversed palindromic sequences, whereas the dimers constrained at the N-terminus bind the same sequence with less helical contents. The stability of specific binding complexes was not affected by the differences in the chirality of the template. However, the stability of the half-specific complex was dramatically affected by the particular enantiomer of the template.

Introductions

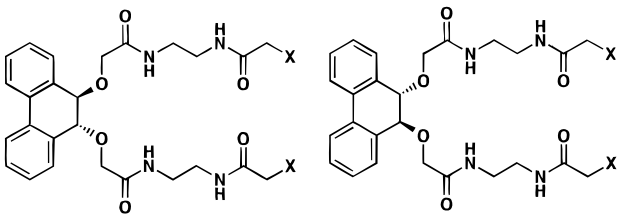
The basic region peptides derived from the basic leucine zipper proteins (bZIP)[1] afford ideal systems for studying the sequence-specific DNA binding of short peptide dimers since the disulfide-bonded dimer of the basic region peptides is sufficient for the sequence-specific DNA binding.[2,3] It is now possible to design peptide dimers that possess DNA binding specificities different from those of the native bZIP by appropriately arranging two basic region peptides with an artificial dimerization domain such as a bulky metal complex,[4−6] a disulfide linkage,[7−9] a host−guest inclusion complex of β-cyclodextrin and the adamantyl group,[10−12] enantiomeric bridged biphenyl derivatives,[13,14] and a lysine residue.[15] With a bulky metal complex as a dimerization domain for the basic region peptide of GCN4, the dimeric peptide preferentially binds the CRE (5′-ATGACGTCAT-3′) sequence over the non-palindromic AP1 (5′-ATGACTCAT-3′) sequence.[4−6] Because the native GCN4 dimer prefers the AP1 sequence, switching the dimerization module from the leucine-zipper coiled coil into the metal complex changes the sequence selectivity of the basic region dimer of GCN4. Similarly, preferential binding to the CRE sequence over the AP1 sequence was observed when a host−guest inclusion complex of β-cyclodextrin and the adamantyl group was used as a non-covalent dimerization module for the basic region peptide of GCN4.[10] It has been shown that an N-terminal disulfide dimer of the basic region of v-Jun binds to a palindromic DNA sequence in which an arrangement of each half site is reversed in terms of the polarity.[7] These studies clearly demonstrate that it is possible to design novel peptide dimers and oligomers by appropriately arranging the DNA contacting unit. However, an underlying question in the design of DNA binding peptide dimers is how to increase the selectivity of these peptide dimers. Effects of the tethering position and the shape of the covalent dimerization module of the peptide dimers on the sequence-selective DNA binding of peptide dimers remain to be answered.

We are interested in how the three-dimensional shape of the dimerization module contributes to the sequence-specific DNA recognition by protein dimers and approach this question through the arrangement of monomeric peptide motifs on a C_2-symmetric chiral template[16] that acts as an artificial dimerization module (Figure 1). We have already tethered the DNA contacting region of MyoD to the enantiomeric templates either at the C-terminus or N-terminus to yield four different constrained peptide dimers.[13] It has been demonstrated that only one of

* Address correspondence to these authors at Kyoto University.
† Kyoto Institute of Technology.
⊗ Abstract published in *Advance ACS Abstracts,* April 1, 1997.

(1) Hurst, H. *Protein Profile* **1994**, *1*, 123.
(2) Talanian, R. V.; McKnight, C. J.; Kim, P. S. *Science* **1990**, *249*, 769.
(3) Talanian, R. V.; McKnight, C. J.; Rutkowski, R.; Kim, P. S. *Biochemistry* **1992**, *31*, 6871.
(4) Cuenoud, B.; Schepartz, A. *Science* **1993**, *259*, 510.
(5) Cuenoud, B.; Schepartz, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1154.
(6) Palmer, C. R.; Sloan, L. S.; Adrian, J. C., Jr.; Cuenoud, B.; Paolella, D. N.; Schepartz, A. *J. Am. Chem. Soc.* **1995**, *117*, 8899.
(7) Park, C.; Campbell, J. L.; Goddard, W. A., III *Proc. Natl. Acad. U.S.A.* **1992**, *89*, 9094.
(8) Park, C.; Campbell, J. L.; Goddard, W. A., III *Proc. Natl. Acad. U.S.A.* **1993**, *90*, 4892.
(9) Park, C.; Campbell, J. L.; Goddard, W. A., III *J. Am. Chem. Soc.* **1995**, *117*, 6287.
(10) Ueno, M.; Murakami, A.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1993**, *115*, 12575.
(11) Ueno, M.; Sawada, M.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1994**, *116*, 11137.
(12) Morii, T.; Yamane, J.; Aizawa, Y.; Makino, K.; Sugiura, Y. *J. Am. Chem. Soc.* **1996**, *118*, 10011.
(13) Morii, T.; Shimomura, M.; Morimoto, M.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 1150.
(14) Okagami, M.; Ueno, M.; Makino, K.; Shimomura, M.; Saito, I.; Morii, T.; Sugiura, Y. *Bioorg. Med. Chem.* **1995**, *3*, 777.
(15) Pellegrini, M.; Ebright, R. H. *J. Am. Chem. Soc.* **1996**, *118*, 5831.
(16) Morii, T.; Murakami, A.; Makino, K.; Morimoto, S.; Saito, I. *Tetrahedron Lett.* **1994**, *35*, 1219.

S0002-7863(96)03280-5 CCC: $14.00



RDHP (1a): X = I
NR (1b): X = C/G23
N6R (1c): X = C6/G23
CR (1d): X = G23/C

SDHP (2a): X = I
NS (2b): X = C/G23
N6S (2c): X = C6/G23
CS (2d): X = G23/C

G23/C: Ac-DPAALKRARNTEAARRSRARLKLQ$\mathbb{C}$-NH_2
C/G23: Ac-$\mathbb{C}$PAALKRARNTEAARRSRARLKLQ-NH_2
C6/G23: Ac-DPAAL$\mathbb{C}$RARNTEAARRSRARLKLQ-NH_2

Figure 1. Structures of the differently constrained GCN4 basic region dimers NR, NS, N6R, and N6S. NR and N6R possess a right-handed geometry for two peptides, whereas NS and N6S possess a left-handed geometry. Coupling of the peptides to the enantiomeric templates was performed by reaction between the thiol group of Cys (C) and the iodoacetyl groups on the templates **1a** and **1b** as described.[13,14]

these four peptide dimers binds specifically to the DNA sequence that is recognized by native MyoD. Specifically, both right- and left-handed C-terminus dimers bind the DNA sequence recognized by the native MyoD. However, competition experiments with a nonspecific competitor DNA have revealed that the peptide dimer with right-handed and C-terminus constraints binds more specifically to the native MyoD binding sequence. The C-terminus dimers of the GCN4 basic region peptide, CR and CS, bind AP1 (5′-ATGACTCAT-3′) and CRE (5′-ATGACGTCAT-3′) DNA sequences that are recognized by the native GCN4 in the same sequence-specific DNA recognition mechanisms, but stabilities of CR− and CS−CRE complexes are quite similar in this case.[14]

We report here factors that increase the sensitivity of sequence-specific DNA binding of geometrically constrained peptide dimers. Dimers of the GCN4 basic region peptide that are constrained at the 6th position from the N-terminus recognized novel palindromic DNA sequences in which the polarity of each half-site of the parent GCN4 binding DNA sequence is reversed. This is in contrast with dimers that are constrained at the N-terminus which failed to recognize the reversed DNA sequences. The stability of specific binding complexes was not affected by the differences in the chirality of the enantiomeric template. However, the stability of half-specific DNA binding complexes of the peptide dimers was affected by the particular enantiomer of the template.

Results

Synthesis of the Basic Region Peptide Dimers with Non-Native Configuration. The C_2-symmetric templates *R*-DHP (**1a**) and *S*-DHP (**1b**) derived from (9*R*,10*R*)- and (9*S*,10*S*)-*trans*-9,10-dihydrophenanthrene-9,10-diol[16] were used as the covalent dimerization module to constrain the relative orientation of two basic region peptides. The amino acid sequences of oligopeptides C/G23 and C6/G23 are derived from that of the basic region of yeast transcription factor GCN4[17] (Figure 1). Unique cysteine residue was incorporated at the N-terminal (C/G23) or at the 6th position (C6/G23) of each peptide. Coupling reactions of the peptides to the enantiomeric templates were performed by reactions between the thiol group of Cys and the iodoacetyl groups on the templates as previously described.[13] Four different constrained dimers have been obtained by using two enantiomeric templates (*R*- and *S*-DHP) and two peptides (C/G23 and C6/G23). NS and NR are constrained at the N-terminus of C/G23 peptide, while N6R and N6S are at the 6th position of C6/G23.

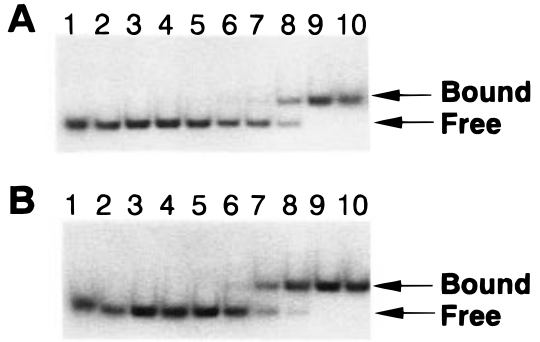


Figure 2. Gel mobility shift assay[18,19] showing the titration with increasing amounts of N6R (A) and N6S (B) to ^{32}P-end-labeled R2CRE21. No peptide was added to the reaction in lane 1. Peptide concentrations in lanes 2−10 respectively were the following: 2, 4, 10, 20, 40, 100, 200, 400, and 2000 pM. Binding reactions were carried out as described in the Experimental Section.

DNA Binding of N−N Dimers Studied by Gel Shift Assay. Two DNA sequences RCRE (5′-TCATCGATGA-3′) and R2CRE (5′-GTCATATGAC-3′) were designed to possess a reversed half-site orientation as compared to the native GCN4 binding sequence CRE (5′-ATGACGTCAT-3′). In the RCRE sequence, the central CG sequence of the CRE remains unchanged, while orientation of the outer 4 bp sequences has been reversed. In the case of R2CRE, orientation of the whole half-site of CRE, 5′-ATGAC-3′, is reversed. The 5′-ATGA-3′ is regarded as a half-site and the central CG step as a spacer sequence for the RCRE, and 5′-ATGAC-3′ is regarded as a half-site for the R2CRE sequence.

DNA binding of each dimer was studied with oligonucleotides containing the RCRE and R2CRE sequences (RCRE21 and R2CRE21) by gel mobility shift assay[18,19] (20 mM Tris−HCl, 4 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, pH 7.5) at 4 °C. N6R and N6S gave mobility-shifted bands corresponding to the 1:1 dimer−DNA complexes with RCRE and R2CRE (Figure 2, parts A and B). Mobilities of these 1:1 dimer−DNA complex bands are almost identical with that for the 1:1 CR− and CS−CRE complexes.[14] No mobility-shifted band corresponding to 2:1 or higher dimer−DNA complexes was observed. Dissociation constants of the complexes between N6R or N6S and the target sequences were obtained by fitting the experimentally obtained titration data to eq 4 (Figure 3, parts A and B). N6R and N6S bind the R2CRE sequence with dissociation constants of 0.15 and 0.10 nM, respectively. Dissociation constants of N6R−RCRE and N6S−RCRE complexes are 0.22 and 0.33 nM, respectively. The results of direct titration indicate that N6R and N6S bind to RCRE and R2CRE sequences with similar affinity.

The C/G23 dimers NR and NS bind RCRE or R2CRE less efficiently than the C6/G23 dimers. The mobility-shifted bands indicate formation of a 1:1 NR−R2CRE complex (Figure 4A). However, no saturation of the binding was obtained for the gel shift titration. Moreover, the bands corresponding to the free DNA become smeary in the presence of higher concentrations of NR (lanes 5 to 10, Figure 4A), and the bound band shows retarded mobility in the presence of 200 nM NR (lane 10). This is also the case for NS with R2CRE (Figure 4B). Thus difference in the tethering position between two peptides resulting in an altered binding behavior of peptide dimers.

Sequence-Specific DNA Binding of N−N Dimers Studied by DNaseI Footprinting. The sequence-specific DNA binding

(17) Hinnebush, A. G. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6442.
(18) Gamer, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047.
(19) Freid, M.; Crothers, D. M. *Nucleic Acids Res.* **1981**, *9*, 6505.

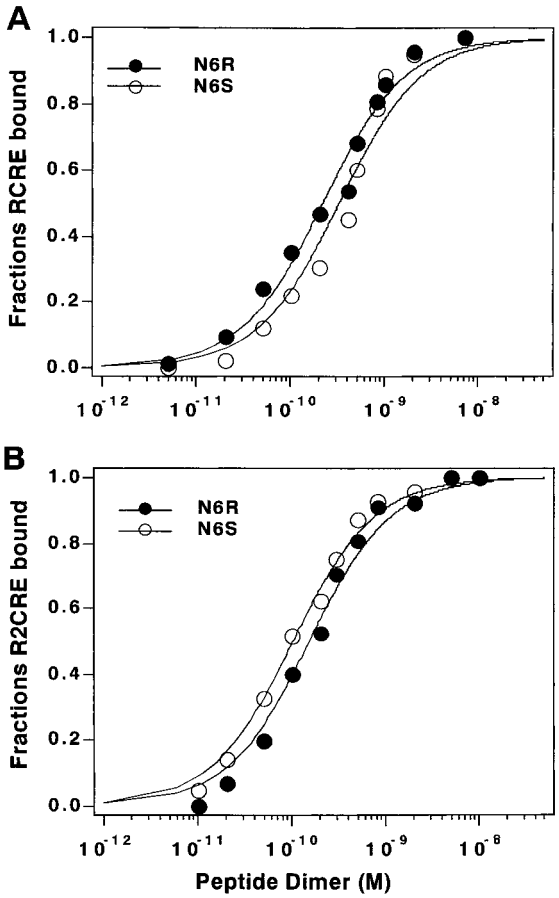


Figure 3. Gel shift titrations of N6R and N6S to RCRE21 (A) and R2CRE21 (B) oligonucleotides. Semilogarithmic plots show the fraction of ^{32}P-labeled DNA bound to N6R (filled circles) and N6S (open circles) as a function of added peptide dimer. Each data point is the average value obtained from three titration experiments. The solid curves are the best fit binding titration isotherm obtained from a nonlinear least-squares algorithm to eq 4.

of these N−N dimers was analyzed by DNase I footprinting[20] with a singly 5′-^{32}P-labeled 46-mer oligonucleotide duplex containing the R2CRE sequence (20 mM Tris−HCl, 4 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, pH 7.5, 4 °C). Distinct footprinting patterns were observed at the R2CRE sequence in the binding complexes of N6R (Figure 5, lanes 4 to 6) and N6S (lanes 7 to 9) dimers. NR (lanes 10 to 12) and NS (lanes 13 to 15) dimers also marked footprints at the R2CRE sequence, but the size of the footprints is larger than that obtained with N6R or N6S. The footprint extends outward from the core R2CRE sequence for both 5′- and 3′-directions in the case of NR and NS (lanes 12 and 15).

Structures of the Basic Region Peptides Bound to DNA. The basic region of GCN4 has disordered structure in the absence of a specific DNA sequence, but it is structured to an α-helix upon binding to a specific DNA sequence.[1,21−25] By using this characteristic of the basic region peptide, structures of the N6R, N6S, NR, and NS bound at the R2CRE sequence were analyzed by using circular dichroism (CD) spectroscopy in the absence or presence of DNA. Spectra of the peptide

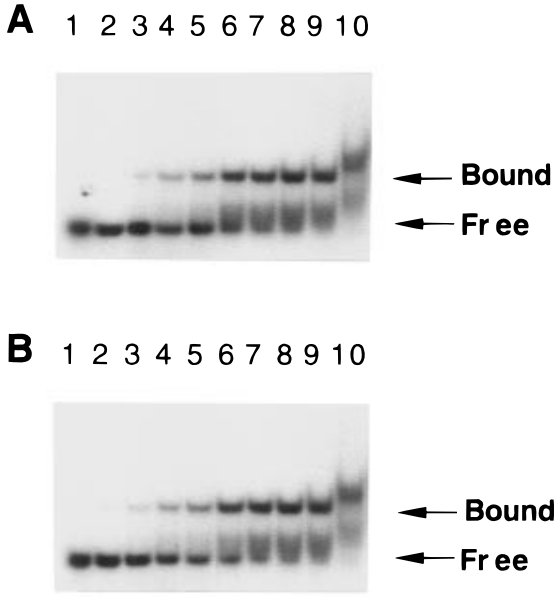


Figure 4. Gel mobility shift assay[18,19] showing the titration with increasing amounts of NR (A) and NS (B) to ^{32}P-end-labeled R2CRE21. No peptide was added to the reaction in lane 1. Peptide concentrations in lanes 2−10 respectively were the following: 0.04, 0.08, 0.2, 0.4, 2, 4, 8, 20, and 80 nM. Binding reactions were carried out as described in the Experimental Section.

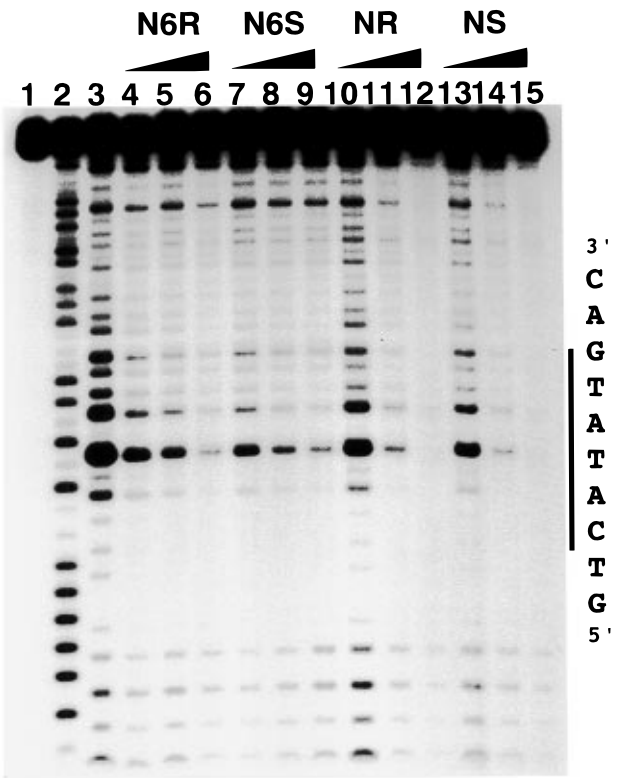


Figure 5. The DNase I footprinting[20] pattern of the R2CRE sequence by NR, NS, N6R, and N6S indicates specific binding of N6R and N6S at the R2CRE sequence. The R2CRE sequence was indicated on the right side of the autoradiogram. Lane 1, intact DNA; lane 2, G + A chemical reaction marker; lane 3, DNase I standard; lanes 4−6, 50, 100, and 200 nM N6R, respectively; lanes 7−9, 50, 100, and 200 nM N6S, respectively; lanes 10−12, 50, 100, and 200 nM NR, respectively; lanes 13−15, 50, 100, and 200 nM NS, respectively.

dimers in the presence of DNA were calculated as the difference between the bound spectrum and free DNA spectrum.[2,4,10,11] Because the enantiomeric templates *R*- and *S*-DHP exhibit CD signals below 240 nm,[16] the difference CD spectra of N- and N6-dimers do not resemble the typical CD spectrum for the

(20) Gallas, D. J.; Schmitz, A. *Nucleic Acids Res.* **1978**, *5*, 3157.
(21) Weiss, M. A.; Ellenberger, T. E.; Wobble, C. R.; Lee, J. P.; Harrison, J. C.; Struhl, K. *Nature* **1990**, *347*, 575.
(22) O'Neil, K. T.; Schuman, J. D.; Ampe, C.; DeGrado, W. F. *Biochemistry* **1991**, *30*, 9030.
(23) Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223.
(24) König, P.; Richmond, T. J. *J. Mol. Biol.* **1993**, *233*, 139.
(25) Keller, W.; König, P.; Richmond, T. J. *J. Mol. Biol.* **1995**, *254*, 657.

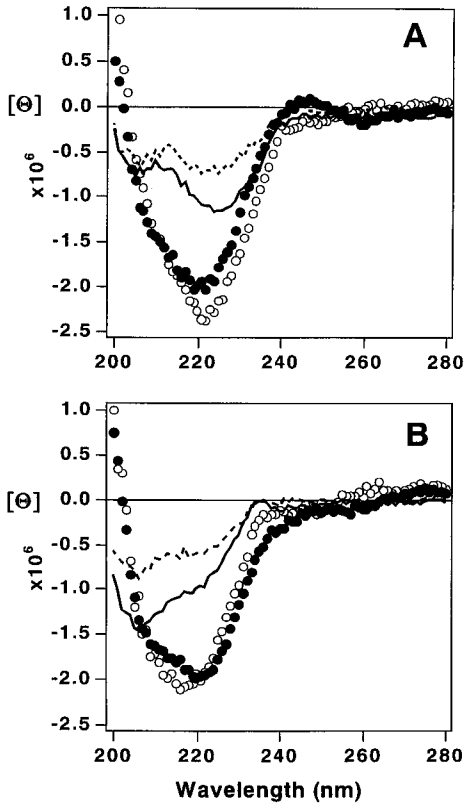


Figure 6. Circular dichroism difference spectra for N6R (A) and N6S (B) show the helical conformation of peptides in the presence of R2CRE sequence. Spectra in the absence of DNA (solid line) and in the presence of NON (dotted line) and R2CRE21 (filled circles), and CR (A) and CS (B) in the presence of CRE21 (open circles) at 4 °C. Spectra of the dimers in the presence of oligonucleotides were calculated as described in the Experimental Section.

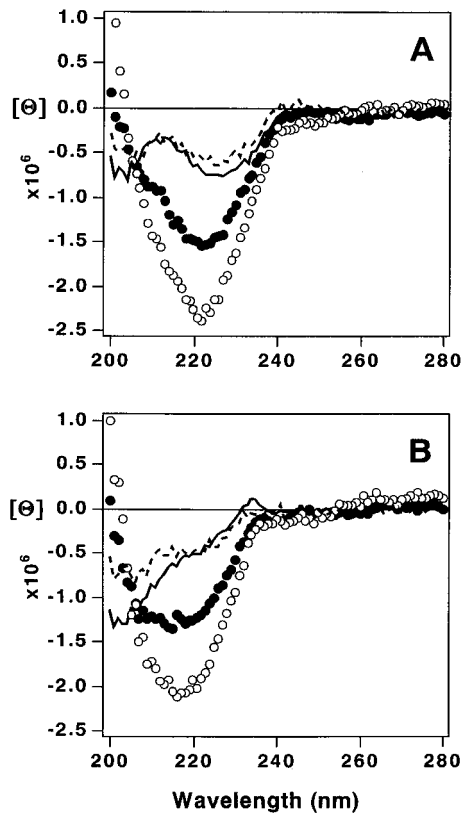


Figure 7. Circular dichroism difference spectra for NR (A) and NS (B) indicate that the helical contents of NR and NS are smaller than those of N6R and N6S. Spectra in the absence of DNA (solid line) and in the presence of NON (dotted line) and R2CRE21 (filled circles), and CR (A) and CS (B) in the presence of CRE21 (open circles) at 4 °C. Spectra of the dimers in the presence of oligonucleotides were calculated as described in the Experimental Section.

helical peptide (double minima at 208 and 222 nm). However, the intensity of the negative CD signal at 222 nm was increased significantly on addition of R2CRE to N6R (Figure 6A). Such a CD spectral change usually reflects the formation of well-ordered secondary structures, such as an α-helical structure, of the peptide. Importantly, the spectrum of N6R in the presence of R2CRE is indistinguishable from the spectrum obtained for the C-terminal basic region dimer (CR, **1d**) in the presence of the CRE sequence. Since CR has been shown to bind the CRE sequence in the helical structure,[14] it is quite likely that N6R also binds the R2CRE sequence in the helical conformation. N6S also revealed an increase of the helical band in the presence of R2CRE (Figure 6B). As was the case for N6R, the spectrum of N6S in the presence of R2CRE was quite similar to that of CS in the presence of CRE, and CS has also been shown to bind the CRE sequence in the helical structure.[14] Both N6R and N6S dimers exhibited a decrease in the intensity of helical band on addition of an oligonucleotide with the noncognate sequence (NON). Thus observed CD spectral changes of N6R and N6S are sequence specific.

NR and NS also exhibited the increase in intensity of CD signals around 222 nm in the presence of R2CRE (Figure 7, parts A and B). Comparison of the CD spectrum of NR in the presence of R2CRE and that of CR in the presence of CRE reveals that the intensity of the CD signal around 222 nm of NR bound to R2CRE is about half that of CR in the presence of CRE (Figure 7A). This is also the case for the intensity of the CD signals around 222 nm of NS as compared to that of CS in the presence of CRE (Figure 7B). In addition, intensities of the CD signals around 222 nm of NR and NS in the presence of R2CRE21 are lower than those of N6R and N6S in the presence of R2CRE21, respectively.

Sequence-Selective Recognition of DNA by N6R and N6S. Sequence selectivity of the DNA binding of N6R and N6S was studied by the gel-shift competition experiment[18,19] with oligonucleotides CRE21, AP20, HS21, and non-related sequences NON and ^{32}P-end-labeled R2CRE21 (Figure 8, parts A, B, C, and D). Dissociation constants of the complexes between the dimers and these oligonucleotides were obtained by fitting the experimentally obtained titration data to eq 10. Affinities of N6R to the noncognate sequences, especially to the "half-matched" sequences, differ considerably from that of N6S (Table 1). The N6R−CRE21 complex is more stable than the N6S−CRE21 complex by a factor of 25. The N6R−AP20 complex is more stable than the corresponding N6S complex by a factor of 11. Similarly, the N6R−HS21 complex is 8 times more stable than N6S−HS21. In the case of non-relating sequence NON, the N6R complex is more stable than the N6S complex by a factor of 3.

Discussion

Design of the GCN4 Basic Region Peptide Dimers with Non-Native Monomer Arrangements and Their Target DNA Sequences. GCN4 is known to bind DNA as a homodimer with each basic region directly contacting the major groove of the half-site DNA.[23−25] The leucine-zipper coiled coil constrains at the C-termini of two GCN4 basic regions. Our modeling study based on the X-ray structures of GCN4 bound to AP1[23] and CRE[24,25] sequences has indicated that a distance between the two arms of the template might be shorter than the ideal distance between two N-terminal positions of the peptides when the N-terminus dimer (NR and NS) bound a palindromic sequence (RCRE and R2CRE) in which the polarity of each

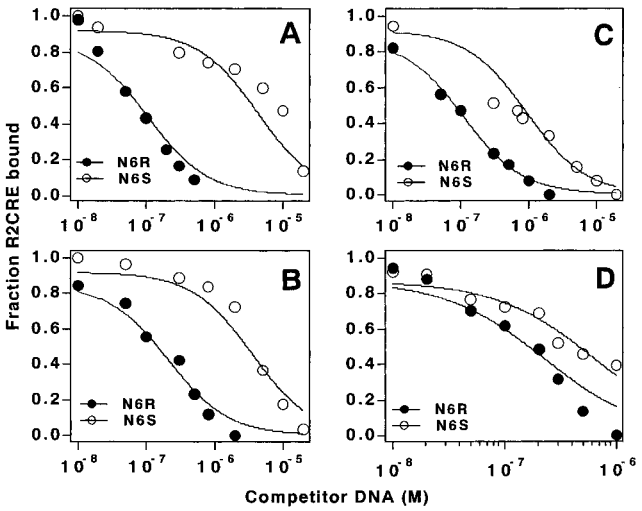


Figure 8. Gel shift competition analysis[18,19] of the relative affinities of N6R and N6S to CRE21 (A), GRE21 (B), HS21 (C), and NON (D) oligonucleotides. Semilogarithmic plots show the fraction of ^{32}P-labeled R2CRE21 bound to N6R (filled circles) and N6S (open circles) as a function of added competitor DNA. Each data point is the average value obtained from three titration experiments. The solid curves are the best fit binding titration isotherm obtained from a nonlinear least-squares algorithm for eq 10. Binding reactions were carried out in the presence of N6R or N6S (1 nM) and indicated amounts of competitor oligonucleotides with 20 pM 5′-^{32}P-labeled oligonucleotide in a binding mixture containing 20 mM Tris−HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 10 μg/mL acetylated BSA, 0.1% NP-40, and 5% glycerol and analyzed as described in the Experimental Section.

Table 1. Equilibrium Dissociation Constants (K_d) Obtained for Binding Complexes of N6R and N6S Dimers to RCRE21, R2CRE21, CRE21, GRE20, HS21, and NON

	K_d (nM)	
	N6R	N6S
RCRE21	0.22[a]	0.33[a]
R2CRE21	0.15[a]	0.10[a]
CRE21	14[b]	363[b]
GRE20	27[b]	300[b]
HS21	15[b]	79[b]
NON	29[b]	83[b]

[a] Obtained from direct titration of the gel shift at 4 °C. [b] Obtained from the gel shift competition assay at 4 °C.

half-site is reversed from that of the CRE sequence. Thus, a cystein residue was replaced for the 6th lysine residue of the GCN4 basic region in the case of C6/G23. The mutated lysine residue at the 6th position is not included in any critical contact to DNA in the X-ray structures of GCN4-DNA complexes,[23−25] and a distance between the lysine residue of each monomer seems to be ideal for accommodating the template **1a** or **1b**. Comparison of the DNA binding of the C/G23 peptide dimers (NR and NS) to that of the C6/G23 peptide dimers (N6R and N6S) would reveal how the spacing between two DNA contacting regions affects the sequence-specific DNA binding of peptide dimers, while comparison of the binding of *R*-template dimers to that of the *S*-template dimers would reveal the effects of a subtle structural change on the dimerization module in the sequence-specific DNA recognition.

The basic region peptides of N−N dimers NR (**1b**), N6R (**1c**), NS (**2b**), and N6S (**2c**) are arranged in reversed polarity as compared to the C−C dimers, such as the native GCN4 dimer and synthetic CR (**1d**) and CS (**2d**).[14] Because the single GCN4 basic region peptide correlates with a half-site of the palindromic DNA sequence (5′-ATGAC-3′), we would expect that the N−N dimer recognizes a palindromic DNA sequence in which each half-site is arranged in a reversed polarity as compared to the CRE sequence. In fact, it has been reported that a disulfide-linked N−N dimer of the basic region peptide derived from v-Jun binds the 5′-TCATCGATGA-3′ sequence.[7]

Binding of N−N Dimers to Palindromic Sequences. Results of the gel mobility shift assay indicate that N6R and N6S form 1:1 binding complexes with the R2CRE (Figure 2) and RCRE (data not shown) sequences. As was observed for the complexes of CR and CS with the CRE sequence,[14] N6R and N6S showed similar affinities to RCRE or R2CRE despite the difference in structure of the dimerization module (Figure 3). Stabilities and mobilities of these binding complexes are similar to those of CR− or CS−CRE21 complexes (K_d = 0.25 nM).[14] Specific formation of the N6R− and N6S−R2CRE complexes is characterized by DNase I footprinting as shown in Figure 5. The size of the footprints marked at the R2CRE sequence with N6R or N6S is comparable to that obtained for CRE−CR or CRE−CS complexes. These results clearly indicate the R2CRE−N6R and R2CRE−N6S complexes are formed in a sequence-specific manner.

Although both NR and NS afforded mobility-shifted bands corresponding to 1:1 complexes with R2CRE (Figure 4, parts A and B), neither NR nor NS binds to a full saturation to R2CRE. This is also the case with the RCRE sequence. NR (Figure 5, lanes 10 to 12) and NS (Figure 5, lanes 13 to 15) gave footprints at the R2CRE sequence, but the size of the footprints at the R2CRE sequence marked by NR or NS is larger than those obtained by N6R or N6S. The amino acid sequences of C/G23 and C6/G23 are identical except for the 6th position, and the essential amino acid residues that contact to the nucleic acid bases or phosphate groups in the co-crystal of GCN4 and its cognate DNA sequence remain unchanged. Because C/G23 dimers (NR and NS) and C6/G23 dimers (N6R and N6S) differ only in the tethering position, the difference in the sequence-specific binding must result from the difference in the dimerization position.

A possible binding model that explains the footprinting results for NR and NS is that one of the basic region peptides of the dimer specifically binds to the half-site of the palindromic R2CRE sequence in the helical conformation and the other peptide of the dimer fits the other half-site without forming any significant secondary structure. Such an extended conformation of the peptide would result in the larger footprints as observed in Figure 5.

Structures of the Basic Region Peptides on DNA. Because the chiral templates *R*- and *S*-DHP have CD signals below 240 nm,[16] it is difficult to compare the peptide structures of *R*- and *S*-dimers quantitatively. However, the difference CD spectra of CR and CS in the presence of CRE21 can serve as the reference spectra of the specific binding complexes, because both CR and CS have been shown to bind the CRE sequence in the helical conformation.[14] If the peptides of N6R and N6S fit in the major groove of RCRE and R2CRE sequences in the helical structure, the difference CD spectra of the specific binding complexes N6R− and N6S−R2CRE would be essentially identical to that of the CR− and CS−CRE complexes, respectively. This is born out for N6R and N6S in the presence of R2CRE21 (Figure 6, parts A and B). These CD data and the footprinting results strongly demonstrate that N6R and N6S specifically bind to RCRE and R2CRE sequences predominantly in the helical conformation.

A plausible explanation for the DNase I footprinting results is the half-specific binding mode for the complexes of NR and NS with the R2CRE sequence. Although other explanations, such as non-specific binding, are possible for the larger footprints obtained with NR and NS, the difference CD spectra of NR and NS in the presence of R2CRE21 also support this

notion. Intensities of the helical band (222 nm) of NR and NS increase upon addition of R2CRE21 (Figure 7, parts A and B), but not to the same extent as observed for N6R and N6S. Such an increase in the intensities of the helical band was not observed in the presence of non-specific sequence NON. It is likely that only one of the C/G23 peptides of NR or NS fits to the half-site of R2CRE sequence in the helical structure.

DNA Binding Specificities of the N6R and N6S Dimers. As observed in the direct titration experiments, N6R− and N6S−R2CRE complexes exhibit similar equilibrium dissociation constants, 0.15 and 0.10 nM, respectively. Interestingly, N6R and N6S showed different affinities to the sequences containing only one half-site, 5′-ATGAC-3′ (Figure 8, parts A, B, and C). The CRE sequence contains two half-sites, 5′-ATGAC-3′, but the polarity of two half-sites for the CRE sequence is opposite from that of the R2CRE sequence. If one of the basic region peptides of N6R or N6S binds to the half-site of the CRE sequence, another basic region peptide will no longer fit to another ATGAC half-site due to the geometrical constraints of the dimers. The same is true for the N6R and N6S complexes with the GRE and HS sequences. Thus given the fact that N6R and N6S form specific complexes with the RCRE or R2CRE sequence, N6R and N6S would form the half-specific complexes with CRE, GRE, and HS. Because the observed dissociation constant of the N6S−CRE complex is larger than that of the N6R−CRE complex, the geometrical constraints by SDHP lower the stabilities of half-specific binding complexes of N6S. On the other hand, the stability of the N6R−NON complex is similar to that of the N6R complexes with CRE, GRE, and HS (Table 1). Such differences in the stabilities of non-specific binding complexes that are affected by a particular enantiomer of the template have also been observed for the MyoD derived peptide dimers.[13] Although the mechanism for this antiselectivity is yet to be established, the chiral template of N6S *reduces the affinity of the peptide dimer for the non-targeting sequence*, rather than increasing the affinity for the specific sequences, possibly by interacting with the N-terminal portions of the peptides. Such an effect of the synthetic dimerization module has also been suggested for the DNA binding of metallopeptides.[6]

Conclusions. We have studied the DNA binding of covalently-bonded peptide homodimers in which the geometrical constraints are systematically varied. The constraint that controls the polarity of the two peptides determines the polarity of the palindromic DNA sequence. The position of covalent bond constraints on the peptides, that is, the spacing between the two peptides, is critical for the specific complex formation. One effect of the chiral templates is a marked reduction in the affinity of the peptide dimer for non-targeting DNA sequences. Since the bZIP basic region dimers already have high affinities for their respective target DNA sequences, the strategy of reducing the nonspecific binding would aid in the design of the second generation homo- and heterodimers that recognize a variety of DNA sequences with high selectivity.

Experimental Section

Materials. *R,R*- and *S,S*-DHP (**1a** and **2a**) were synthesized as described previously.[13,14] Pentafluorophenyl esters of protected Fmoc (9-fluorenylmethoxycarbonyl) amino acid were obtained from Novabiochem. 1-Hydroxybenzotriazole (HOBt) was from Nakaraitesque. PAL resin (0.36 mmol/g) was from PerSeptive Biosystems. Dimethylformamide was dried over CaH_2, distilled from ninhydrine at 55 °C under reduced pressure, and stored over molecular sieves 4A. Protected nucleoside phosphoramidites were from PerSeptive Biosystems. T4 polynucleotide kinase was obtained from New England Biolab. DNase I was from Takara. Calf thymus DNA was from Pharmacia. [γ-^{32}P]-ATP was from Amersham. BSA was from Clonetech. HPLC grade acetonitrile (Nakaraitesque) was used for both analytical and preparative HPLC. A reagent grade Milli-Q water was used throughout the experiments. Gel electrophoresis grade acrylamide, bisacrylamide, and urea were obtained from Wako Chemicals. All other chemicals were reagent grade and were used without further purification. Sephadex G-10 and G-25 were obtained from Pharmacia. A reverse-phase C18 column (20 × 250 mm, Ultron VX-Peptide, Sinwa Chemical Industry) was used for purification of peptides and the peptide dimers for preparative purposes. Analytical HPLC was carried out on a reverse-phase C18 column (4.6 × 150 mm, Ultron VX-Peptide, Sinwa Chemical Industry). Oligonucleotides were purified on a reverse-phase C18 column (6 × 150 mm, Ultron VX-Nucleotide, Sinwa Chemical Industry). Amino acid analyses were performed with an AccQ Tag Chemistry Package (Waters) according to a company protocol. Proton NMR spectra were recorded at 500 MHz on a Brüker ARX500 spectrometer. Chemical shifts are represented in parts per million relative to residual HOD. Electrospray mass spectra were recorded in the positive ion mode on a Perkin-Elmer Sciex API III.

Synthesis of GCN4 Basic Region Peptides. Oligopeptides C/G23 and C6/G23 were prepared from a single stepwise manual solid-phase peptide synthesis with pentafluorophenyl esters of Fmoc (9-fluorenylmethoxycarbonyl) amino acids (Novabiochem) and 1-hydroxybenzotriazole.[26] The coupling reaction was performed with 0.3 g of PAL resin (0.36 mmol equiv/g) and Fmoc amino acid pentafluorophenyl ester (0.25 mmol) in the presence of 1-hydroxybenzotriazole (0.25 mmol) in anhydrous DMF for 45 min. Completion of the coupling was monitored by Kaiser test,[27] and the coupling reaction was repeated until completion. Removal of the Fmoc group was performed by treatment with 20% piperidine−DMF. The amino termini were acetylated with acetylimidazole, and the peptides were cleaved from the resin with a cleavage mixture containing bromotrimethylsilane (1.35 mL), thioanisole (1.2 mL), 1,2-ethanedithiol (0.6 mL), and *m*-cresol (0.2 mL) in trifluoroacetic acid (7.48 mL), then desalted by Sephadex G-10 chromatography in 5% acetic acid. Purification was done by reverse-phase HPLC with a Ultron VX-Peptide column (Sinwa Chemical Industry; 20 × 250 mm) and a linear gradient of acetonitrile−water with 0.2% trifluoroacetic acid (8 to 30% acetonitrile in 50 min; flow rate 5 mL/min).

Reaction of Chiral Templates with Oligopeptides C/G23 and C6/G23. A dimethylformamide solution of (9*R*,10*R*)-DHP (3.7 mg, 5 nmol) was added to an aqueous solution (0.1 M $NaHCO_3$, pH 8.0) containing oligopeptide C/G23 (15 nmol) at 0 °C under nitrogen atmosphere. The resulting solution was kept at 0 °C for 3 h, then allowed to warm up to room temperature and quenched with addition of acetic acid. Purification of the crude product by Sephadex G-25 (5% acetic acid as an eluent) and reversed-phase HPLC on a Ultron VX-Peptide column (Sinwa Chemical Industry; 20 × 250 mm) with 0.2% trifluoroacetic acid−acetonitrile as an eluent yielded the disulfide-bonded dimer $(C/G23)_2$, disubstituted (9*R*,10*R*)-$(C/G23)_2$DHP (NR), and monosubstituted (9*R*,10*R*)-(C/G23)DHP (order of appearance in the chromatogram). Syntheses of NS, N6R, and N6S were carried out in a similar manner. Amino acid analysis (cysteine was not determined): NR, expected $D_2T_2S_2E_4A_{12}L_4K_4R_{12}P_2$, found $D_{2.03}T_{1.90}S_{1.82}E_{4.00}A_{12.30}L_{4.06}K_{3.98}R_{12.17}P_{1.92}$; NS, expected $D_2T_2S_2E_4A_{12}L_4K_4R_{12}P_2$, found $D_{2.03}T_{1.90}S_{1.82}E_{4.00}A_{12.35}L_{4.06}K_{3.98}R_{12.18}P_{1.93}$; N6R, expected $D_4T_2S_2E_4A_{12}L_4K_2R_{12}P_2$, found $D_{4.00}T_{1.89}S_{1.82}E_{3.93}A_{12.16}L_{4.01}K_{1.92}R_{11.85}P_{1.90}$; N6S, expected $D_4T_2S_2E_4A_{12}L_4K_2R_{12}P_2$, found $D_{4.00}T_{1.90}S_{1.81}E_{3.96}A_{12.20}L_{4.02}K_{1.95}R_{12.02}P_{1.92}$. MS (ion spray; 50% acetonitrile, 0.05% formic acid): NR, calculated for $[M^+]$ 5822.81, found 5822.52; NS, calculated for $[M^+]$ 5822.81, found 5821.89; N6R, calculated for $[M^+]$ 5796.64, found 55795.66; N6S, calculated for $[M^+]$ 5796.64, found 5795.20.

Synthesis and 5′-End Labeling of Oligonucleotides (CRE21, GRE20, HS21, RCRE21, R2CRE21, NON, and R2CRE46). Oligonucleotides were synthesized on a PerSeptive Biosystems DNA synthesizer with a standard method and purified by reverse-phase HPLC on an Ultron VX-Nucleotide column (Sinwa Chemical Industry; 6 × 150 mm) with 0.1 M triethylammonium acetate−acetonitrile as an eluent. The oligonucleotides were labeled by kinase reaction with use

(26) Atherton, E.; Sheppard, R. C. *J. Chem. Soc., Chem. Commun.* **1985**, 165.

(27) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595.

of [γ-^{32}P]ATP (5000 Ci/mmol) and T4 polynucleotide kinase,[28] followed by removal of unincorporated label over Sep-pak (Millipore Waters). The oligonucleotide was then denatured and annealed to a 2-fold molar excess of the opposite strand in 10 mM Tris−HCl, 50 mM NaCl at pH 8.0. Nucleotide sequences of the oligonucleotides used in the present study are RCRE21 (5′-CGGTCATCGATGATTTTTTTC-3′), R2CRE21 (5′-CGGGTCATATGACTTTTTTTC-3′), CRE21 (5′-CGGATGACGTCATTTTTTTTC-3′), GRE20 (5′-CGGATGACTCATTTTTTTTC-3′), HS21 (5′-CGGATGACACTGCTTTTTTTC-3′, NON (5′-GATCCCCCCAACACCTGCTGCCTGA-3′), and R2CRE46 (5′-CCGAATTCCACAGTGAAAAGTCATATGACCCGGGCAATGGATCCGC-3′. Sequences of only a single strand are shown.

Gel Mobility Shift Assay.[18,19] Binding reactions were carried out in the presence of the indicated peptide dimer with ∼20 pM 5′-^{32}P-labeled oligonucleotide (double-stranded) in a binding mixture containing 20 mM Tris−HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 10 μg/mL BSA, 0.5% NP-40, and 5% glycerol. The binding mixtures were incubated at 4 °C for 30 min, and an aliquot (5 μL) of each binding mixture was directly loaded onto an 8% nondenaturing acrylamide gel (29:1 acrylamide/bis-acrylamide), run in TBE buffer (20 mM Tris, 20 mM boric acid, and 0.4 mM EDTA) at 4 °C, and analyzed by autoradiography. Increasing concentration of the peptide dimer was used for the direct titration of band-shift. Increasing concentrations of non-radio-labeled oligonucleotide (double-stranded) was added to the binding mixture for the band-shift competition assay. The increase or decrease of the mobility-shifted band was quantitated by the densitometry of the autoradiogram by using NIH image (version 1.58). Concentrations of the peptide dimers were determined with an extinction coefficient at 270 nm (ϵ_{270}) of 17 000 M^{-1} cm^{-1}, and were confirmed by quantitative amino acid analyses with α-aminobutyric acid as an internal standard (Waters AccQ-Tag Chemistry Package, Millipore).

DNase I Footprinting.[20] A 46-mer double-stranded oligonucleotide R2CRE46 was prepared as described above. The reaction mixture contained 20 mM Tris−HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 20000 cpm singly 5′-^{32}P-end-labeled 46-mer DNA, and peptide dimer where indicated, in 20 μL total volume. Nuclease digestion was initiated on addition of 1 unit of DNase I on ice and was quenched by addition of 20 μL of 0.6 M NaOAc, 10 mM EDTA, and 0.4 mM calf thymus DNA after 30 s. Samples were purified by phenol−chloroform extraction and ethanol precipitation, suspended in 80% formamide loading dye, run on a 12% sequencing acrylamide gel and analyzed by autoradiography.

Determination of Dissociation Constants. Method A. Direct Measurement of Dissociation Constants. Using an assumption that the peptide dimer bound with 1:1 stoichiometry to the single recognition site present in each DNA, the binding may be represented by eqs 1 and 2.

$$\mathrm{PD} \overset{K_d}{\rightleftharpoons} \mathrm{P} + \mathrm{D} \tag{1}$$

$$K_d = \frac{(\mathrm{P})(\mathrm{D})}{\mathrm{PD}} \tag{2}$$

where P represents the peptide dimer, D represents the DNA fragment, PD represents the peptide dimer−DNA complex, and K_d is the dissociation constant that governs the reaction. The ratio of bound to total DNA, represented as θ, can be written as

$$\theta = \frac{\mathrm{PD}}{(\mathrm{P}) + (\mathrm{D})} \tag{3}$$

Substituting PD from eq 2 gives

$$\theta = \frac{P_T}{(K_d + P_T)} \tag{4}$$

where P_T is the total concentration of the peptide dimer. The K_d was obtained by fitting the experimentally obtained binding data θ to the theoretical eq 4 with a nonlinear least-squares fitting program (Igor 2.02; WaveMetrics Inc., Lake Oswego OR).

Method B. Dissociation Constants Obtained from Gel Retardation Competition Assay. Using an assumption that the peptide dimer bound with 1:1 stoichiometry to the single recognition site present in each DNA, the binding of the peptide dimer to radio-labeled and competitor DNA may be represented by eqs 5−8.

$$\mathrm{PD} \overset{K_{PD}}{\rightleftharpoons} \mathrm{P} + \mathrm{D} \tag{5}$$

$$\mathrm{PC} \overset{K_{PC}}{\rightleftharpoons} \mathrm{P} + \mathrm{C} \tag{6}$$

$$K_{PD} = \frac{(D_T - \mathrm{PD})(P_T - \mathrm{PD} - \mathrm{PC})}{\mathrm{PD}} \tag{7}$$

$$K_{PC} = \frac{(C_T - \mathrm{PD})(P_T - \mathrm{PD} - \mathrm{PC})}{\mathrm{PC}} \tag{8}$$

P_T, D_T, and C_T represent the total concentrations of the peptide dimer, radio labeled DNA, and competitor DNA, respectively. PD and PC are the concentrations of peptide−radiolabeled DNA complex and peptide−competitor DNA complex, respectively. K_{PD} represents K_d of the peptide−radiolabeled DNA complex, and K_{PC} represents the K_d of the peptide−competitor DNA complex. The fractional saturation of the radiolabeled DNA with the peptide dimer may be represented by eq 3. Thus the fractional saturation θ for the competition reaction is represented by eq 9

$$\theta = \frac{P_T(1-\theta)}{K_{PD}\left(1 + \frac{C_T}{K_{PC}}\right) + D_T(1-\theta)} \tag{9}$$

By solving the quadratic expression of eq 9, θ can be written as

$$\theta = \left\{K_{PD} + \frac{K_{PD}}{K_{PC}}C_T + P_T + D_T - \left[\left(K_{PD} + \frac{K_{PD}}{K_{PC}}C_T + P_T + D_T\right)^2 - 4D_TP_T\right]^{1/2}\right\} \Big/ 2D_T \tag{10}$$

The K_{PC} for each competitor DNA was obtained by fitting the fractional saturation data θ obtained experimentally to the theoretical eq 10 with a nonlinear least-squares fitting program (Igor 2.02; WaveMetrics Inc., Lake Oswego OR). The dissociation constants for the radiolabeled DNA (K_{PD}) were obtained by the direct titration method. A typical value for D_T is 20 pM. The K_{PD} obtained by varying K_{PD} and K_{PC} simultaneously agrees well with those determined by Method A.

Measurement of CD Spectra. Spectra of the peptide dimers in the presence of oligonucleotides were calculated as the difference between the bound spectrum and a spectrum of the respective free oligonucleotide. CD spectra were obtained with a Jasco J-720 CD spectrometer at 4 °C in a 1-mm cell. Samples contained 20 mM Tris−HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 4 μM peptide dimer, and 5 μM oligonucleotide duplex when present. Spectra were the average of 32 scans and were corrected with a spectrum of buffer alone but not smoothed.

Acknowledgment. This work was supported by the Grant-in-Aid for Scientific Research on Priority Areas (No. 08219222 and No. 06240104) from the Ministry of Education, Science, Sports and Culture, Japan.

JA963280K

(28) Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1987.