Pergamon

Bioorganic & Medicinal Chemistry, Vol. 3, No. 6, pp. 777–784, 1995

0968-0896/95 $9.50 + .00

0968-0896(95)00047-X

Sequence-Specific DNA Binding by Covalently Constrained Peptide Dimers of the Basic Leucine Zipper Protein GCN4

Masako Okagami,[a] Masaru Ueno,[a] Keisuke Makino,[a] Masatoshi Shimomura,[b] Isao Saito,[b] Takashi Morii*[c] and Yukio Sugiura[c]

[a]*Department of Polymer Science and Engineering, Kyoto Institute of Technology, Sakyo-ku, Kyoto 606, Japan*
[b]*Department of Synthetic Chemistry, Kyoto University, Sakyo-ku, Kyoto 606, Japan*
[c]*Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan*

Abstract—DNA binding of covalently bonded peptide dimers was studied by using enantiomeric and C_2-symmetric templates as a dimerization module. Amino acid sequence of the peptide is derived from that of DNA contact region of the basic leucine zipper protein GCN4. These peptide dimers were designed to possess different constraints with respect to the orientation of two peptides. The basic region peptides were covalently linked to the enantiomeric template at the C-terminal ends. Two peptides are arranged either in a right-handed or left-handed geometry depending on the chirality of the template. The GCN4 basic region dimers with both right-handed and left-handed geometries show equal affinity to the native GCN4 binding DNA sequences, 5'-ATGACTCAT-3' and 5'-ATGACGTCAT-3', as revealed by the gel mobility shift assay. Specific recognition of the palindromic DNA sequence by the peptide dimers was confirmed by the DNase I footprinting. Circular dichroism spectroscopic study indicates that the basic region peptides bound the target DNA sequence in a helical conformation. The degree to which a chiral constraint effects may depend on the geometry of two DNA binding domains in the parent protein–DNA complex and on a position to apply the chiral constraint.

Introduction

One of the fundamental questions in understanding gene regulation is the mechanism by which transcription factors recognize particular DNA sequences. Despite recent progress in structural and biochemical studies of the sequence-specific DNA binding proteins, a universal code for the recognition between proteins and nucleic acids has yet to be generalized.[1] However, these studies have revealed several features of how the proteins recognize specific DNA sequences. For example, the proteins use relatively small regions to directly contact several base pairs of DNA, and most of the sequence-specific DNA binding proteins become functional in dimeric or oligomeric forms. The first characteristic is well demonstrated by the discovery of growing numbers of DNA-binding motifs, such as helix–turn–helix,[1,2] homeo-domains,[1,3] basic-leucine zippers (bZIP),[1,4] basic-helix–loop–helix (bHLH),[1,5] and at least three types of zinc fingers.[1,6] These motifs tend to use α-helices in the direct-contact to the DNA major groove. Because most of the direct contacts between the amino acids and nucleic acid bases are made within such recognition helices, the rest of the proteins can be regarded as an architecture to position the recognition helices at a proper geometry with respect to the specific DNA sequences.[1b] Formation of a well ordered dimer would determine the relative orientation of each monomer. Thus, positioning of the recognition helices is further constrained by such a quaternary structure formation. The shape and size of the dimerization module would be critical for the final positionings of the recognition helices.

We are interested in the question of how the three-dimensional shape of the dimerization module could contribute to sequence-specific DNA recognition by protein dimers.[7] In the simplest case, the direct DNA-reading head can be regarded as a simple α-helix. There are at least two types of constraints possible for arranging two α-helices upon formation of a dyad symmetric protein dimer. One is a constraint of the polarity that allows two orientations, either in an N-terminal to N-terminal (N–N) or in a C-terminal to C-terminal (C–C) arrangement for two helices. Another one is a chiral constraint that defines a right-handed or a left-handed arrangement of two helices with respect to the C_2 axis perpendicular to the major groove of DNA. Combination of these constraints results in four different types of dimers (Fig. 1). Our approach to the evaluation of the role of the three-dimensional constraints on dimer formation uses enantiomeric templates as a covalently bonded dimerization module.[7] By using the enantiomers of C_2 chiral template and tethering either at the C-terminal or N-terminal of peptides, four differently constrained peptide dimers are obtained. We have used this approach to evaluate the DNA binding of peptide dimers derived from the basic region of bHLH protein, MyoD.[7] It has been demonstrated that only one of the four differently constrained peptide dimers binds specifically to a DNA sequence that is recognized by native MyoD.

Specifically, both right- and left-handed C–C dimers bind the DNA sequence recognized by the native MyoD. However, competition experiments with the non-specific competitor DNA have revealed that the peptide dimer with right-handed and C–C constraints binds more specifically to the native MyoD binding sequence. In fact, a recent report on a crystal structure of MyoD bHLH domain–DNA complex proved that the basic region of the MyoD dimer lies in the major groove of DNA in the C–C orientation and right-handed geometry.[8]

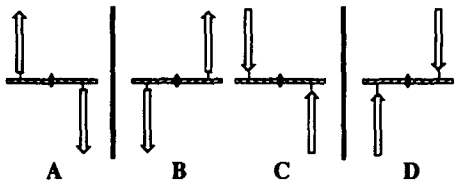

Figure 1. Schematic diagram showing the left-handed (A and C) and right-handed (B and D) arrangements and N-terminal to N-terminal (A and B) and C-terminal to C-terminal (C and D) arrangements of geometrically constrained peptide dimers. Arrows represent the peptides with N-terminal to C-terminal polarity.

In order to study the role of chiral constraint of protein dimers for sequence-specific DNA binding further, it is desirable to use a peptide derived from other structurally well characterized DNA binding proteins. The basic-leucine zipper (bZIP) motif contains a DNA contact region rich in basic amino acid residues (basic region) and a dimerization element (leucine zipper domain).[1,4,9] Dimerization is mediated by a parallel coiled-coil formation of the leucine zipper domain, and the basic region contains sufficient information for specific DNA binding.[10] Several synthetic DNA-binding peptides have been reported by using the basic region of bZIP protein. These include GCN4 basic region peptide dimers with disulfide linkages,[10] with a metal complex as a dimerization module,[11] and with a non-covalent dimerization module.[12] DNA binding of disulfide-bonded basic region dimers with reversed polarity[13a] and with tandem configuration[13b] has also been reported. Because most of the dimer models are constrained only by the distance between two monomers, it would be interesting to compare our system by using the same bZIP type DNA binding peptides. We report here the synthesis and DNA binding studies of geometrically constrained peptide dimers derived from the basic region of a yeast transcriptional activator GCN4. Peptide dimers with different three-dimensional shapes are obtained by changing the chirality on the covalent dimerization module. The peptide dimers with native-like configuration bind the DNA sequences that are recognized by GCN4. Unlike the case of the geometrically constrained dimers with the DNA binding region of MyoD, the peptide dimer with either right-handed or left-handed geometry binds equally to the GCN4 binding sequences. Thus, the effect of the chiral constraint for the DNA binding of peptide dimers may vary depending on the three-dimensional structures of the parent protein–DNA complexes and on a position to apply the chiral constraint.

Experimental

Materials

R,R- and *S,S*-DHP (**1a** and **1b**) were synthesized as described previously.[7,14] 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^{-1}) were from Milligen. Dimethylformamide was dried over CaH_2, distilled from ninhydrine at 55 °C under reduced pressure, and stored over molecular sieves 4 Å. Protected nucleoside phosphoramidites were from Milligen. T4 polynucleotide kinase was obtained from Toyobo. DNase I was from Takara. Poly(dI–dC) 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 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 were 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). Proton NMR spectra were recorded at 500 MHz on a Bruker ARX500 in D_2O. Chemical shifts are represented in parts per million relative to residual HOD.

Synthesis of oligopeptide G23/C

Oligopeptide G23/C was prepared from a single stepwise manual solid-phase peptide synthesis using pentafluorophenyl esters of Fmoc (9-fluorenylmethoxycarbonyl) amino acids (Novabiochem) and 1-hydroxybenzotriazole.[15] Coupling reaction was performed with 0.3 g of PAL resin (0.36 mmol equiv. g^{-1}) 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 the Kaiser test,[16] 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 acetyl-imidazole 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 an Ultron VX-Peptide column (Sinwa Chemical Industry; 20 × 250 mm) and linear gradient of acetonitrile–water with 0.2% trifluoroacetic acid (8 to

30% acetonitrile in 50 min; flow rate: 5 mL min^{-1}). Amino acid sequence of G23/C is, Ac-DPAALKRARN-TEAARRSRARKLQC-NH_2.

[Abbreviations for the amino acids are: A, Ala; C, Cys; D, Asp; E, Glu; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr.]

*Reaction of chiral templates (**1a** and **1b**) with oligopeptide G23/C*

A dimethylformamide solution (200 μL) of **1a** (3.7 mg, 5 nmol) was added to an aqueous solution (0.1 M $NaHCO_3$, pH 8.0, 800 μL) containing oligopeptide G23/C (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 reverse-phase HPLC on an Ultron VX-Peptide column (Sinwa Chemical Industry; 20 × 250 mm) with 0.2% trifluoroacetic acid–acetonitrile as an eluent yielded the disulfide-bonded dimer $(G23/C)_2$, di-substituted CR (**2a**) and mono-substituted **1a** (order of appearance in the chromatogram). Reaction of **1b** with the peptide was carried out in a similar manner. CR (**2a**) ^1H NMR (D_2O) δ 0.89 (*m*, 24H), 1.22 (*d*, 6H, *J* = 6.2 Hz), 1.43 (*m*, 44H), 1.64 (*m*, 36H), 1.81 (*m*, 42H), 2.06 (*m*, 16H), 2.02 (*s*, 6H), 2.35 (*m*, 2H), 2.39 (*t*, 4H, *J* = 7.6 Hz), 2.51 (*t*, 4H, *J* = 7.7 Hz), 2.90 (*m*, 16H), 3.19 (*m*, 36H), 3.83 (*m*, 8H), 4.01 (*d*, 2H, *J* = 16.1 Hz), 4.15 (*d*, 2H, *J* = 15.4 Hz), 4.30 (*m*, 46H), 4.48 (*dd*, 2H, *J* = 7.2 Hz, 5.4 Hz), 4.75 (*s*, 2H), 4.95 (*t*, 2H, *J* = 6.9 Hz), 7.45 (*t*, 2H, *J* = 7.4 Hz), 7.53 (*d*, 2H, *J* = 7.3 Hz), 7.58 (*t*, 2H, *J* = 7.2 Hz), 8.02 (*d*, 2H, *J* = 7.3 Hz). Amino acid analysis (cysteine was not determined); expected: $D_4T_2S_2E_4A_{12}L_4K_4R_{12}P_2$, found: $D_{4.00}T_{1.92}S_{1.85}E_{3.99}A_{12.28}L_{4.08}$-$K_{3.99}R_{12.14}P_{1.89}$. CS (**2b**) ^1H NMR (D_2O) δ 4.01 (*d*, 2H, *J* = 15.8 Hz), 4.13 (*d*, 2H, *J* = 15.8 Hz), 4.75 (*s*, 2H), 7.41 (*t*, 2H, *J* = 7.5 Hz), 7.52 (*d*, 2H, *J* = 7.2 Hz), 7.59 (*t*, 2H, *J* = 7.3 Hz), 8.00 (*d*, 2H, *J* = 7.3 Hz) (only signals derived from the template are listed). Amino acid analysis (cysteine was not determined); expected: $D_4T_2S_2E_4A_{12}L_4K_4R_{12}P_2$, found: $D_{4.00}T_{1.91}S_{1.85}E_{3.98}A_{12.34}L_{4.12}$-$K_{4.00}R_{12.18}P_{1.92}$.

Synthesis and 5'-end labeling of oligonucleotides GRE20, CRE21 and CRE46

Oligonucleotides were synthesized on a Milligen DNA synthesizer with 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[17] using [γ-^{32}P]ATP (5000 Ci $mmol^{-1}$) and T4 polynucleotide kinase, 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 25 mM Tris–HCl, 100 mM NaCl, 1 mM EDTA at pH 8.0. Nucleotide sequences of the oligonucleotides used in the present study are CRE21 (5'-CGGATGACGTCATTTTTTTTC-3'), GRE20 (5'-CGGATGACTCATTTTTTTTC-3'), and CRE46 (5'-TCGAATTCCACAGTGAGAAATGACGTCATCCAGA CTTAGGATCCGC-3': top strand). Sequences of only single strand are shown.

Gel mobility shift assay

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 $MgCl_2$, 2 mM EDTA, 10 μg mL^{-1} BSA, 0.1% NP-40 and 5% glycerol. The binding mixtures were incubated at 4 °C for 30 min, and an aliquot (8 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 0.1 mM EDTA) at 4 °C, and analyzed by autoradiography. Increasing concentration of the peptide dimer was used for the direct titration of band-shift. The increase of the mobility-shifted band was quantitated by the densitometry of the autoradiogram. Concentration of the peptide dimer was determined with an extinction coefficient at 270 nm (e_{270}) of 17,000 M^{-1} cm^{-1},[7] and was confirmed by quantitative amino acid analysis with α-aminobutyric acid as an internal standard (Waters AccQ-Tag Chemistry Package, Millipore).

Determination of dissociation constants

Using an assumption that the peptide dimer binds with 1:1 stoichiometry to the single recognition site present in each DNA, the binding may be represented by equations 1 and 2,

$$PD \underset{}{\overset{K_d}{\rightleftharpoons}} P + D \qquad (1)$$

$$K_d = \frac{PD}{(P)(D)} . \qquad (2)$$

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:

$$\theta = \frac{PD}{PD + D} \qquad (3)$$

Substituting PD from equation 2 gives:

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

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 equation 4 with a nonlinear least-squares fitting program (Igor 1.25; WaveMetrics Inc., Lake Oswego OR).

DNase I footprinting

46-mer Double-stranded oligonucleotides, CRE46 were prepared as described above. Reaction mixture contained 20 mM Tris–HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 20 μM poly(dI–dC), ~20000 cpm singly 5'-^{32}P-end-labeled CRE46, 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.

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.

Results and Discussion

Synthesis of geometrically constrained peptide dimers

The dimerization domain of protein dimers, the helix–loop–helix module in the previous study[7] and the leucine zipper module in the present study, was replaced with C_2-symmetric templates **1a** and **1b** to synthesize geometrically constrained peptide dimers. These chiral templates are synthesized from *(9R,10R)*- and *(9S,10S)-trans*-9,10-dihydrophenanthrene-9,10-diol as previously reported.[7,14] Assuming that the biphenyl chromophore faces toward the groove of DNA, rather than being exposed to a solvent,[14] the *(9R,10R)*-isomer would achieve a right-handed geometry of two DNA binding peptides, while (9*S*,10*S*)-isomer would achieve a left-handed geometry. The iodoacetyl group of **1a** and **1b** was used to generate a covalent linkage between the peptide and each chiral template through a specific reaction with a SH group of unique cysteine residue in the peptides.

The amino acid sequence of 24-mer peptide G23/C is derived from that of the DNA contacting region (basic region) of the yeast transcriptional activator GCN4.[9] Unique cysteine residue is incorporated at the C-terminal of the peptide. G23/C was synthesized as C-terminal amide and the N-terminal was acetylated. GCN4 is known to bind DNA as a homodimer with each basic region directly contacting the major groove of DNA. The dimerization is mediated through a coiled-coil structure at the leucine-zipper domain that is located at the C-terminal of the basic region. Since the three-dimensional structures of the complexes of GCN4 and its cognate DNA binding sites are available,[18] using the basic region of GCN4 for our system would be ideal to investigate an effect of different constraint induced by the chiral templates. Kim and co-workers have reported a minimum amino acid length required for the sequence-specific DNA binding of the basic region peptide dimer.[10b] According to their results, G23/C has a sufficient length to bind the GCN4 binding sequence at 4 °C when arranged into a dimeric form through a (Gly–Gly–Cys-)$_2$ linkage at the C-terminal of the peptide.

The coupling reaction between the sulfhydryl group of G23/C peptide and the iodoacetyl group of enantiomeric template (**1a** and **1b**) was carried out in an aqueous alkaline solution under nitrogen atmosphere at 0 °C.[7] The crude reaction mixture was purified with gel filtration and reverse-phase HPLC to give two differently constrained peptide dimers CR (**2a**) and CS (**2b**) (Fig. 2). Each dimer possesses different chirality in the arrangement of the G23/C peptide. Specifically, peptide dimer CR is expected to bind in a right-handed geometry, and the dimer CS to bind in a left-handed geometry.

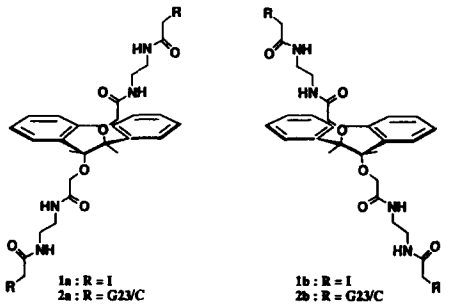

Figure 2. Structures of the GCN4 basic region dimers CR (**2a**) and CS (**2b**).

DNA binding of the basic region peptide dimers CR and CS

It has been demonstrated that the native GCN4 dimer specifically binds 5'-ATGACTCAT-3' (GRE) and 5'-ATGACGTCAT-3' (CRE) sequences in a similar affinity.[19] Sequence-specific DNA binding of CR and CS dimers to the GRE and CRE sequences was compared by gel mobility shift assay.[20] Typical binding titration experiments are shown in Figures 3A and 4A. Equilibrium dissociation constants of the complexes formed between CR or CS and either CRE21 or GRE20 were obtained by fitting a theoretical curve to the gel mobility shift titration data (Figs 3B and 4B).

The results are summarized in Table 1. Apparent dissociation constants of the peptide dimers CR and CS

for the CRE sequence are almost identical. This is also the case for the GRE sequence. Thus, each GCN4 basic region dimer with either the right-handed (CR) or left-handed (CS) geometry shows almost identical affinity to the native GCN4 binding sequences. In the case of MyoD basic region dimers with the same chiral templates, the right-handed dimer showed slightly higher affinity to the native MyoD binding sequence.[7] Possible explanations for the reason why these two systems yielded different results come from an inspection of the X-ray crystal structures of GCN4 and MyoD bound to their cognate DNA sequences.[8,18] Both bHLH and bZIP proteins juxtapose the α-helices rich in basic amino acid residues to contact DNA major groove. However, the way in which two basic regions are positioned in the GCN4 dimer is different from that in the MyoD dimer. The C-terminal portions of two basic regions of the GCN4 dimer are positioned next to each other because of the structural constraint induced by the leucine zipper coiled-coil. On the other hand, the C-terminal portions of two basic regions are separated further in the MyoD dimer. Such differences in the arrangements of DNA contacting regions in the parent protein–DNA complexes could yield different results for the MyoD system and GCN4 system. The C-terminal cysteine of G23/C corresponds to Arg249 in the native

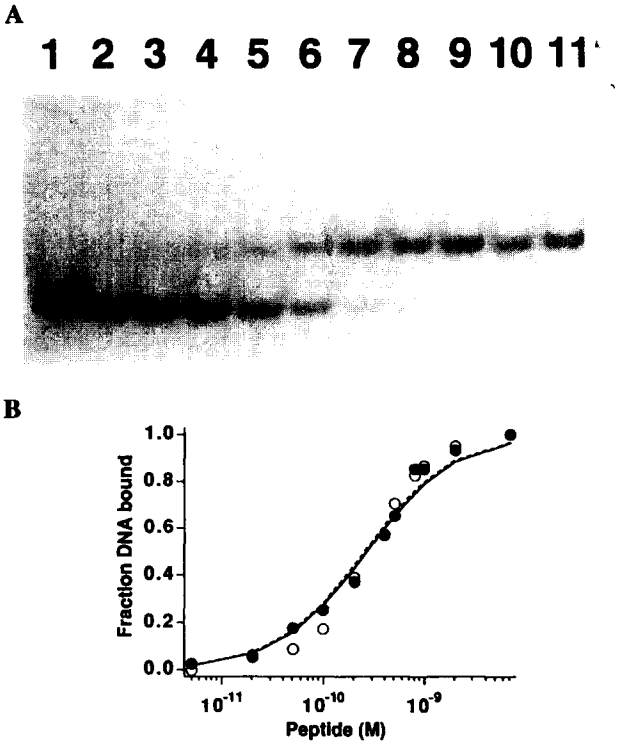

Figure 3. Determination of the dissociation constants for CRE21–CR and CRE21–CS complexes using the gel mobility shift assay. **A.** Titration with increasing amounts of CR. Each reaction contained ~20 pM ^{32}P end-labeled CRE21 DNA. No CR was added to the reaction in lane 1. CR concentrations in lanes 2–10, respectively, were: 0, 5, 20, 50, 200, 400, 500, 800, 1000, 2000 and 7000 pM. **B.** Quantitation of data for CRE21–CR and CRE21–CS complexes. Symbols represent average data from panel A and companion experiments. Open circles represent data for CR and the filled circles for CS. The solid curve is the best fit binding titration isotherm obtained from a nonlinear least-squares algorithm using equation 4 for CR and the dotted curve is for CS.

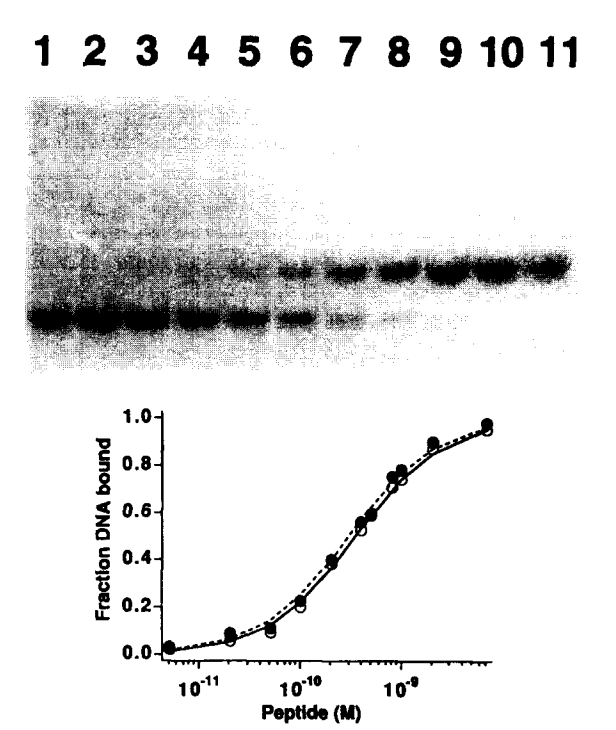

Figure 4. Determination of the dissociation constants for GRE20–CR and GRE20–CS complexes using the gel mobility shift assay. **A.** Titration with increasing amounts of CR. All reactions contained ~20 pM ^{32}P end-labeled GRE20 DNA. No CR was added to the reaction in lane 1. CR concentrations in lanes 2–10, respectively were: 0, 5, 20, 50, 200, 400, 500, 800, 1000, 2000 and 7000 pM. **B.** Quantitation of data for GRE20–CR and GRE20–CS complexes. Symbols represent average data from panel A and companion experiments. Open circles represent data for CR and the filled circles for CS. The solid curve is the best fit binding titration isotherm obtained from a nonlinear least-squares algorithm using equation 4 for CR and the dotted curve is for CS.

Table 1. Dissociation constants obtained for binding complexes of CR and CS dimers to CRE21 and GRE20

	CR (nM)[a]	CS (nM)[a]
CRE21	0.26	0.25
GRE20	0.35	0.30

[a]Obtained by fitting the experimentally obtained binding data at 4 °C to the theoretical equation.

GCN4. It is notable that Arg249 of each monomer locates almost parallel in the complex between GCN4 dimer and DNA. This parallel orientation of the C-terminal cysteine would make CR and CS dimers less sensitive to the effect of chiral template.

Native GCN4 dimer binds both CRE and GRE sequences with comparable affinity.[19a] Although structure of the dimerization domain is apparently different from that of the native GCN4, CR and CS dimers also bind the CRE and GRE sequences with almost identical affinity. Selective binding to the CRE sequence over the GRE sequence is reported for a metal complex mediated dimer of the GCN4 basic region peptide.[11] In this case, the bulky metal complex itself would contribute significantly to the binding site selectivity. The GCN4 basic region dimers with

disulfide linkage, which would not possess any structural constraint other than the distance between two C-terminal ends, showed 2 to 3 fold greater affinity to the CRE sequence than that to the GRE sequence.[10b] It has been demonstrated that six amino acid residues in between the basic region and the leucine zipper, termed spacer segment, contribute to the CRE/GRE selectivity.[21,22,23] The G23/C peptide contains only the first three residues of the spacer segment. Thus, the basic region itself and structure of the enantiomeric template would be responsible for the observed selectivity of CR and CS dimers to the CRE and GRE sequences.

DNase I footprinting study

Sequence-specific DNA binding of CR and CS was analyzed at nucleotide resolution by using DNase I footprinting assay.[24] A 46-mer double-stranded DNA containing unique CRE sequence (CRE46) was used as a binding substrate. Figure 5 shows an autoradiogram of the footprinting study with the 46-mer DNA fragment that was 5'-^{32}P end-labeled either at the top or bottom strand. A distinct footprinting pattern was marked at the CRE sequence in the presence of CR dimer (Fig. 5, lanes 4 to 6 and 13 to 16). Similarly, clear footprints were observed for CS as well (Fig. 5, lanes 1 to 3 and 16 to 18). In addition, the size of footprint is almost equivalent to that observed for the GCN4 dimer.[10,11a,25] The enhanced cleavage near the binding site, which is often observed in the DNase I footprinting experiments of DNA binding molecules, was also observed at the equivalent position to that observed for GCN4. Results from the gel mobility shift assay and the DNase I footprinting experiments clearly indicate that (i) both CR and CS bind the specific DNA sequences that are recognized by the native GCN4 dimer and (ii) the way in which the basic region peptide of CR and CS lies on the major groove of specific DNA sequence is quite similar to that observed for the native GCN4.

Structure of the basic region peptides in the presence of specific DNA sequences

The basic region of GCN4 gets structured in an α-helical conformation upon binding to the specific DNA sequence. Changes in the secondary structure of the peptides could be monitored from a differential CD

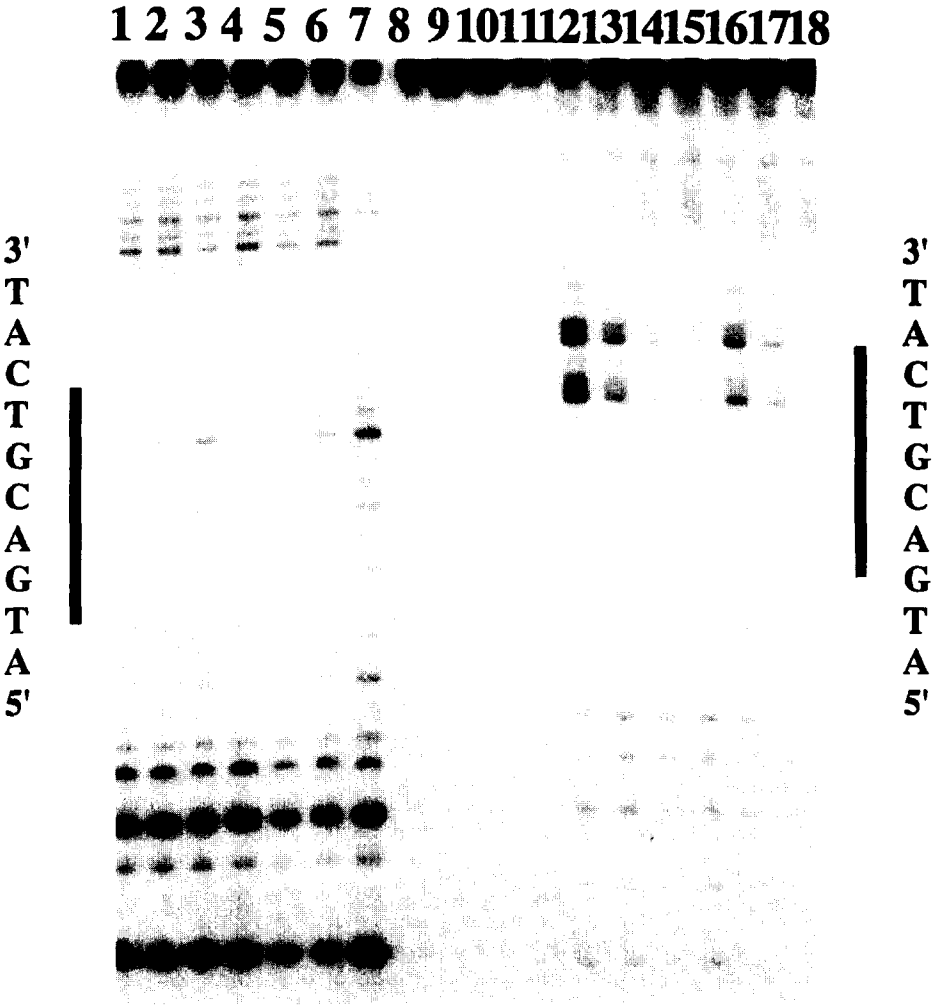

Figure 5. DNase I footprinting experiments with CR and CS on the 5'-^{32}P-labeled CRE46 at the top and at the bottom strand: lanes 1–9, labeled at top strand; lanes 10–18, labeled at bottom strand; lanes 9 and 10, intact DNA; lanes 8 and 11, G + A chemical reaction marker; lanes 7 and 12, DNase I standard; lanes 4–6 and 13–15, 50, 100 and 200 nM CR, respectively; lanes 7–9 and 16–18, 50, 100 and 200 nM CS, respectively. The 5'-ATGACGTCAT-3' binding site is shown on the left and right sides of the autoradiogram.

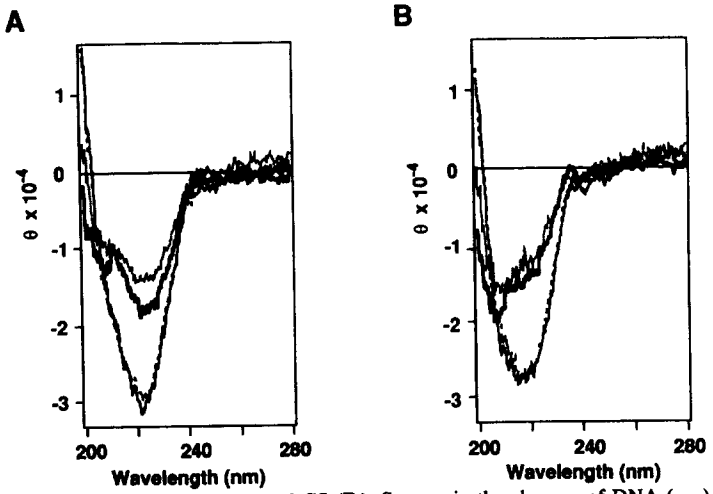

Figure 6. Circular dichroism difference spectra for CR (**A**) and CS (**B**). Spectra in the absence of DNA (—), and in the presence of NON21 (·····), CRE21 (—), and GRE20 (– – –) at 4 °C. Spectra of the dimers in the presence of oligonucleotides were calculated as described in the Experimental section.

spectrum that is obtained by subtracting the CD signal of free DNA from that of the peptide dimer–DNA complex.[10,11,12] The helical structure of peptides is easily characterized by monitoring the CD signal at 222 nm,[27] which is attributable to the helical band. Figures 6A and B show CD spectra of CR and CS, respectively, in the absence or presence of various oligonucleotides. The chiral templates show strong CD bands around 220 nm which reflect the helicity of the bridged-biphenyl chromophore of **1a** and **1b**.[14] The chiral template **1a** exhibits a positive CD signal at 235 nm and a negative signal below 220 nm, while **1b** displays signals with the opposite sign. Although the CD signal from the chiral template interferes with the calculation of an exact helical content of the bound peptide, we can judge the conformational change of the peptide at least in a qualitative manner. The DNA-bending properties of many bZIP proteins, including Fos and Jun, have been reported.[28]However, it has been shown that binding of GCN4 to the CRE sequence does not accompany the DNA-bending.[23a] Because little change was observed in the differential CD spectra above 260 nm (Figs 6A and 6B), it is likely that no major DNA structural distortion occurred upon binding of CR and CS to the CRE sequence. The CD signal at 222 nm changed little in the presence of the oligonucleotide NON. In contrast, intensity of the helical band increased significantly in the presence of a stoichiometric amount of CRE21 for both CR and CS. The helical band increased only in the presence of the CRE and/or GRE sequences. These results indicate that the G23/C peptide of CR and CS dimers recognize DNA sequences in a similar mechanism as observed for the native GCN4 dimer.

Conclusion

We have synthesized geometrically constrained dimers of a peptide derived from GCN4 by using enantiomeric templates. These peptide dimers show the sequence specific DNA binding identical to that of GCN4. Changing the chirality for the arrangement of the peptide did not show any effect on the DNA binding of the GCN4 basic region peptide dimer. The result is different from that observed for the peptide dimers derived from MyoD. Our results on the DNA binding of the MyoD basic region dimers and GCN4 basic region dimers indicate that an effect of changing the chirality for the arrangement of a pair of peptides is rather small as compared to that of changing the polarity. However, the degree to which the chiral constraint effects may depend on (i) the geometry of two DNA binding domains in the native protein–DNA complex and (ii) a position to apply the chiral constraint.

Acknowledgements

This work was supported by Sasakawa Scientific Research Grant from the Japan Science Society and the Grant-in Aid for Scientific Research on Priority Areas No. 06240232 from the Ministry of Education, Science and Culture, Japan to T.M. M.U. is a research fellow of the Japan Society for the Promotion of Science.

References

1. (a) Harrison, S. C.; Aggarwal, A. K. *Ann. Rev. Biochem.* **1990**, *59*, 933; (b) Pabo, C. O.; Sauer, R. T. *Ann. Rev. Biochem.* **1992**, *61*, 1053.

2. (a) Brennan, R. G.; Matthews, B. W. *J. Biol. Chem.* **1989**, 264, 1903; (b) Gehring, W. J.; Müller, M.; Affolter, M.; Percival-Smith, A.; Billeter, M.; Qian, Y. Q.; Otting, G.; Wüthrich, K. *Trends Genet.* **1990**, *6*, 323.

3. (a) Kissinger, C. R.; Liu, B.; Martin-Blanco, E.; Kornberg, T. B.; Pabo, C. O. *Cell* **1990**, *63*, 579; (b) Wolberger, C.; Vershon, A. K.; Liu, B.; Johnson, A. D.; Pabo, C. O. *Cell* **1991**, *67*, 517; (c) Gehring, W. J.; Qian, Y. Q.; Billerter, M.; Furukubo-Tokunaga, K.; Schier, A. F.; Resendez-Perez, D.; Affolter, M.; Otting, G.; Wüthrich, K. *Cell* **1994**, *78*, 211.

4. (a) Landschultz, W. H.; Johnson, P. F.; McKnight, S. L. *Science* **1988**, *240*, 1759; (b) Vinson, C. R.; Sigler, P. B.;

McKnight, S. L. *Science* **1989**, *246*, 911; (c) O'Neil, K. T.; Hoess, R. H.; DeGrado, W. F. *Science* **1990**, 249, 774.

5. (a) Prendergast, G. C.; Ziff, E. B. *Nature* **1989**, *341*, 392; (b) Murre, C.; McCaw, P. S.; Baltimore, D. *Cell* **1989**, *56*, 777; (c) Lassar, A. B.; Buskin, J. N.; Lockson, D.; Davis, R. L.; Apone, S.; Hauschka, S. D.; Weintraub, H. *Cell* **1989**, *58*, 823; (d) Davis, R. L.; Cheng, P.-F.; Lassar, A. B.; Weintraub, H. *Cell* **1990**, *60*, 733; (e) Sun, X.-H.; Baltimore, D. *Cell* **1991**, *64*, 459; (f) Anthony-Cahill, S. J.; Benfield, P. A.; Fairman, R.; Wasserman, Z. R.; Brenner, S. L.; Stafford, III W. F.; Altenbach, C.; Hubbell, W.; Degrado, W. F. *Science* **1992**, *255*, 979.

6. (a) Klug, A; Rhodes, D. *Trends Biochem.* **1987**, *12*, 464; (b) Vallee, B. L.; Coleman, J. E.; Auld, D. S. *Proc. Natl Acad. Sci. U.S.A.* **1991**, *88*, 999; (c) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809; (d) Marmorstein, R; Carey, M.; Ptashne, M.; Harrison, S. C. *Nature* **1992**, *356*, 408; (e) Kraulis, P. J.; Raine, A. R. C.; Gadhavi, P. L.; Lane, E. D. *Nature* **1992**, *356*, 448; (f) Baleja, J. P.; Marmorstein, R.; Harrison, S. C.; Wagner, G. *Nature* **1992**, *356*, 450; (g) Pavletich, N. P.; Pabo, C. O. *Science* **1993**, *261*, 1701.

7. Morii, T.; Shimomura, M.; Morimoto, M.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 1150.

8. Ma, P. C. M.; Rould, M. A.; Weintraub, H.; Pabo, C. O. *Cell* **1994**, 77, 451.

9. McKnight, S. L. *Sci. Am.* **1991**, *264*, 54.

10. (a) Talanian, R. V.; McKnight, C. J.; Kim, P. S. *Science* **1990**, *249*, 769; (b) Talanian, R. V.; McKnight, C. J.; Rutkowski, R.; Kim, P. S. *Biochemistry* **1992**, *31*, 6871.

11. (a) Cuenoud, B.; Schepartz, A. *Science* **1993**, *259*, 510; (b) Cuenoud, B.; Schepartz, A. *Proc. Natl Acad. Sci. U.S.A.* **1993**, *90*, 1154.

12. (a) Ueno, M.; Murakami, A.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1993**, *115*, 12575; (b) Ueno, M.; Sawada, M.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1994**, *116*, 11137.

13. (a) Park, C.; Campbell, J. L.; Goddard, III, W. A. *Proc. Natl Acad. Sci. U.S.A.* **1992**, *89*, 9094; (b) Park, C.; Campbell, J. L.; Goddard, III W. A. *Proc. Natl Acad. Sci. U.S.A.*, **1993**, *90*, 4892.

14. Morii, T.; Murakami, A.; Makino, K.; Morimoto, S.; Saito, I. *Tetrahedron Lett.* **1994**, *35*, 1219.

15. Atherton, E.; Sheppard, R. C. *J. Chem. Soc. Chem. Commun.* **1985**, 165.

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

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

18 (a) Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223; (b) Konig, P.; Richmond, T. J. *J. Mol. Biol.* **1993**, *233*, 139.

19. (a) Sellers, J. W.; Vincent, A. C.; Struhl, K. *Mol. Cell. Biol.* **1990**, *10*, 5077; (b) Weiss, M. A.; Ellenberger, T. E.; Wobble, C. R.; Lee, J. P.; Harrison, J. C.; Struhl, K. *Nature* **1990**, *347*, 575.

20. (a) Gamer, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047; (b) Freid, M.; Crothers, D. M. *Nucleic Acids Res.* **1981**, *9*, 6505.

21. Johnson, P. F. *Mol. Cell. Biol.* **1993**, *13*, 6919.

22. Kim, J.; Tzamarias, D.; Ellenberger, T.; Struhl, K.; Harrison, S. C. *Proc. Natl Acad. Sci. U.S.A.* **1993**, *90*, 4513.

23. (a) Paolella, D. N.; Palmer, C. R.; Schepartz, A. *Science* **1994**, *264*, 1130; (b) Metallo, S. J.; Schepartz, A. *Chemistry & Biology* **1994**, *1*, 143.

24. Gallas, D. J.; Schmitz, A. *Nucleic Acids Res.* **1978**, *5*, 3157.

25. Oakly, M.; Dervan, P. B. *Science* **1990**, *248*, 847.

26. Schuman, J. D.; Vinson, C. R.; McKnight, S. L. *Science* **1990**, *269*, 771.

27. Chen, Y.-H.; Yang, J. T.; Chau, K. H. *Biochemistry* **1974**, *13*, 3350; (b) Woody, R. W. *The Peptides* p. 15, Undenfriend, S.; Meienhofer, J.; Hruby, V. J., Eds; Academic Press; New York, 1985.

28. (a) Kerppola, T. K.; Curran, T. *Cell* **1991**, *66*, 317; (b) Kerppola, T. K.; Curran, *T. Science* **1991**, *254*, 1210; (c) Kerppola, T. K.; Curran, T. *Mol. Cell. Biol.* **1993**, *13*, 5479.

(*Received in U.S.A.* 28 *November* 1994; *accepted* 17 *February* 1995)