VOLUME 118, NUMBER 42
OCTOBER 23, 1996

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Cooperative Oligomerization Enhances Sequence-Selective DNA Binding by a Short Peptide

Takashi Morii,*,† Junji Yamane,‡ Yasunori Aizawa,† Keisuke Makino,‡ and Yukio Sugiura*,†

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

Received November 6, 1995. Revised Manuscript Received May 30, 1996[⊗]

Abstract: A series of short peptides derived from the basic region of the basic leucine zipper protein GCN4 were synthesized to study the cooperative DNA binding to direct repeat sequences. A modified lysine residue bearing an adamantyl group at the ϵ-amino group was incorporated at the N-terminal position, and β-cyclodextrin was attached at the C-terminal cysteine residue of the parent basic region peptide. The resulting peptide G2AdCd possesses both host and guest molecules in the same peptide chain. DNA binding of the G2AdCd peptides to the single-, double-, and triple-direct-repeat sequences of the CRE half-site was compared by titration of the gel shift. The G2AdCd peptide did not bind the single CRE half-site, although a peptide lacking the β-cyclodextrin group formed a specific monomer−half-site complex. G2AdCd bound the double-direct-repeat sequence as a dimer in a cooperative manner. Moreover, cooperative formation of a 3:1 G2AdCd−DNA complex was observed for a triple-direct-repeat sequence. No monomer−DNA complex of G2AdCd was observed for the double- or triple-direct-repeat sequence. In the absence of DNA, G2AdCd forms an intramolecular host−guest complex. Formation of this cyclic peptide reduces the affinity of monomeric G2AdCd. The highly selective binding of G2AdCd observed here was accomplished by (i) its cooperative nature of DNA binding and (ii) destabilization of its nonspecific DNA binding complex.

The regulation of gene expression is mediated through specific interactions between DNA binding proteins and regulatory DNA sequences. In such interactions, it is quite rare that a monomeric protein with a single structural module binds a specific DNA sequence. Rather DNA binding proteins often operate as homo- or heterodimers, or as higher oligomers.[1−3] The protein−DNA interaction reveals a cooperative nature in these cases, and the protein−protein interaction plays key roles in enhancing the selectivity of specific DNA binding and in increasing the sensitivity of equilibrium binding to the change in protein concentrations.[4,5] Yet it is very difficult to elucidate the precise nature of cooperativity in the specific protein−DNA interaction by using native proteins. Although details of the stereospecific interaction between the DNA recognition motif and DNA have been demonstrated in the X-ray crystal structures,[2,3] there still remains a question of how the multiple protein complexes can form on the specific long sequence of DNA.[6] In order to explain the highly specific recognition of the DNA site by multiple proteins seen in the native system, it is necessary to understand how the interactions between proteins enhance the specific recognition events. For this purpose, the sequence-specific DNA binding of dimeric protein offers the simplest

* To whom correspondence should be addressed.
† Kyoto University.
‡ Kyoto Institute of Technology.
[⊗] Abstract published in *Advance ACS Abstracts*, October 1, 1996.
(1) Jones, N. *Cell* **1990**, *61*, 9 and references therein.
(2) Harrison, S. C.; Aggarwal, A. K. *Annu. Rev. Biochem.* **1990**, *59*, 933.
(3) Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1992**, *61*, 1053.
(4) Ptashne, M. *A Genetic Switch*; Blackwell Scientific Publications and Cell Press: Palo Alto, CA, 1986.
(5) Adhya, S. *Annu. Rev. Genet.* **1989**, *23*, 227.
(6) Tijan, R; Maniatis, T. *Cell* **1994**, *77*, 5.

S0002-7863(95)03741-3 CCC: $12.00

example for the recognition events including both the protein−DNA and the protein−protein interactions.

The basic leucine zipper (bZIP) family of proteins consist of the simplest module for the sequence-specific recognition of DNA.[7] The protein−protein interaction, namely, the dimerization, is mediated through a coiled-coil structure of the leucine zipper domain, while the region rich in basic amino acid residues (the basic region) located N-terminal to the leucine zipper domain directly contacts the DNA major groove. The finding by Kim and co-workers that disulfide-bonded dimers of the basic region specifically bind the DNA sequence as observed for the native bZIP further encourages the use of the basic region as a DNA binding unit for the design of novel DNA binding peptides.[8,9] By extending this method, several disulfide dimers and trimers of basic region peptides that target novel nonpalindromic DNA sequences have been reported.[10−12] Also sequence-specific DNA binding of covalent basic region dimers using metal complexes[13−15] or bridged biphenyl derivatives[16,17] as the covalent dimerization domain has been reported. These studies demonstrated that the basic region by itself can be used as the DNA binding domain.

We have shown previously that an artificial dimerization module consisting of β-cyclodextrin (Cd) and its guest compound (Ad, adamantyl group) mimics a noncovalent and specific protein−protein interaction.[18,19] A peptide with the adamantyl group efficiently binds a specific DNA sequence in a heterodimeric form in the presence of a peptide with Cd.[18] The target DNA sequence possesses either a palindromic nature or a nonpalindromic nature depending on the contents of the peptides.[19] We now ask whether this strategy could extend to a cooperative DNA binding by a *homooligomer* of short peptides. An oligopeptide derived from the basic region of GCN4 was used as the DNA interaction domain. The host and guest molecules (Cd and Ad) were introduced on the peptide to facilitate a noncovalent and specific interaction between the peptides through the host−guest inclusion complex formation. The peptide without β-Cd binds to a half-site of the native GCN4 binding palindromic sequence without any cooperativity. When both host and guest molecules are incorporated into the same peptide chain, binding of the peptide dimers and trimers to tandemly repeated half-sites becomes cooperative. Interestingly, the peptide with both host and guest molecules shows reduced affinity with the single half-site as compared to the peptide without Cd.

Experimental Section

Materials. Protected Fmoc (9-fluorenylmethoxycarbonyl) amino acids, protected Fmoc amino acid 3-hydroxy-4-oxo-3,4-dihydro-1,2,3-benzotriazine esters (ODhbt ester), Fmoc-OSu ((9-fluorenylmethyl)-succinimidyl carbonate) and PyBOP (bromotrispyrrolidinophosphonium hexafluorophosphate) were obtained from Novabiochem. 1-Hydroxybenzotriazole (HOBt) was from Nakaraitesque. Fmoc-PAL-PEG resin (0.2 mmol/g) was from PerSeptive Biosystems. *N*,*N*-Dimethylformamide was dried over CaH_2, distilled from ninhydrin at 55 °C under reduced pressure, and stored over molecular sieves 4A. Protected nucleoside phosphoramidites were from Milligen. T4 polynucleotide kinase was obtained from New England Biolab. [γ-^{32}P]ATP was from Amersham. HPLC grade acetonitrile (Wako Chemicals) was used for both analytical and preparative HPLC. Reagent grade Milli-Q water was used throughout the experiments. Gel electrophoresis grade acrylamide and bisacrylamide 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 reversed-phase C18 column (20 × 250 mm, Ultron VX-Peptide, Sinwa Chemical Industry) was used for purification of peptides for preparative purposes. Analytical HPLC was carried out on a reversed-phase C18 column (4.6 × 150 mm, Ultron VX-Peptide, Sinwa Chemical Industry). Oligonucleotides were purified on a reversed-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 Bruker ARX500 or at 200 MHz on a Varian VXR200 spectrometer. Chemical shifts are represented in parts per million relative to residual HOD. Fast atom bombardment (FAB) mass spectrometry was recorded in the positive ion mode on a JEOL JMS-DX-303 HF spectrometer. Electrospray mass spectrometry was recorded in the positive ion mode on a Perkin-Elmer Sciex API III.

Synthesis of *N*α-Fmoc-*Nϵ*-(adamantylacetyl)lysine. To a solution containing *N*α-*t*-Boc-lysine ethyl ester (1.2 g, 4.3 mmol), adamantaneacetic acid (0.85 g, 4.4 mol), and *N*-hydroxybenzotriazole (0.90 g, 6.6 mmol) in dry *N*,*N*-dimethylformamide (10 mL) was added DCC (1.11 g, 5.4 mmol) at 0 °C. The reaction mixture was stirred to ambient temperature overnight, and the solvent was removed under reduced pressure. The residue was dissolved in ethyl acetate and filtered to remove urea, and the solvent was washed with 10% $NaHCO_3$. The organic layer was separated and dried over Na_2SO_4. The solvent was removed under reduced pressure, and the resulting residue was purified on a silica gel column (*n*-hexane:ethyl acetate = 1:1). The main fraction (R_f = 0.2) was collected to give *N*α-*t*-Boc-*Nϵ*-(adamantylacetyl)lysine ethyl ester as a colorless gum (1.7 g). An ethanolic solution (20 mL) of the ester (4.76g, 10.6 mmol) was added to 10.6 mL of 1 M LiOH at 0 °C, and stirred for 1 h. After removal of ethanol under reduced pressure, the aqueous solution was extracted with ethyl acetate, the aqueous layer was separated, and the solvent was removed under reduced pressure to give a colorless gum. The residue was dissolved in 90% aqueous TFA, and the solution was stirred for 1 h at ambient temperature. Evaporation of the solvent under reduced pressure gave a TFA salt of *Nϵ*-(adamantylacetyl)lysine as a colorless gum (2.35 g, 5.38 mmol). *Nϵ*-(adamantylacetyl)lysine (1.7 g, 3.9 mmol) was dissolved in 10% $NaHCO_3$ (10 mL), Fmoc-OSu (1.1g, 3.3 mmol) in dioxane (5 mL) was added, and the reaction mixture was stirred for 2 h at ambient temperature. The solution was acidified with 1 M HCl and then extracted twice with ethyl acetate. The organic layer was combined, and evaporation of the solvent under reduced pressure gave a colorless gum. The residue was purified on a silica gel column (ethyl acetate:*n*-hexane = 2:1). The main fraction (R_f = 0.6) was collected, and evaporation of the solvent gave *N*α-Fmoc-*Nϵ*-(adamantylacetyl)-L-lysine as a white foam (1.1 g): ^{1}H NMR ($CDCl_3$) δ 1.36−1.71 (m, 17H), 1.83−1.92 (m, 6H), 3.20−3.26 (m, 2H), 4.20 (t, 1H, J = 7 Hz), 4.33−4.44 (m, 3H), 5.70 (t, 1H, J = 6 Hz), 5.78 (d, 1H, J = 8 Hz), 7.29 (dd, 2H, J = 7 Hz), 7.39 (dd, 2H, J = 6, 7 Hz), 7.59 (dd, 2H, J = 7 Hz), 7.75 (d, 2H, J = 7 Hz); high-resolution mass spectrum (FAB, glycerol matrix) *m/z* 545.3027 [(M + H)$^+$; calcd for $C_{33}H_{41}N_2O_5$, 545.3026].

Synthesis of the GCN4 Basic Region Peptides with an Adamantyl Group. Peptides were manually synthesized by using a solid-phase

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

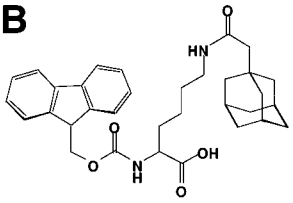
A

G1: Ac-𝕏PAALKRARNTEAARRSRARKLQℂ-NH_2

G2: Ac-D𝕏AALKRARNTEAARRSRARKLQℂ-NH_2

G3: Ac-𝕏DPAALKRARNTEAARRSRARKLQℂ-NH_2

B



Fmoc-Lys(Ad)

C

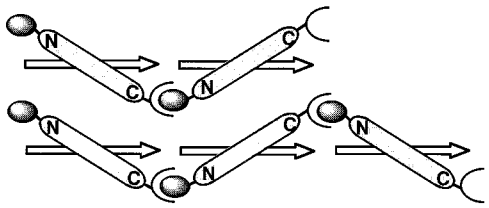


Figure 1. (A) Amino acid sequences for G1, G2, and G3 peptides indicating the locations of *N*ε-(adamantylacetyl)lysine (X). The cysteine residues (C) were modified with 6-deoxy-6-iodo-β-cyclodextrin and iodoacetamide for GAdCd peptides and GAd peptides, respectively. (B) Structure of Fmoc-Lys(Ad). (C) Schematic representations showing the GAdCd homooligomer bound on the direct-repeat sequence of the CRE half-site. Hatched bars represent the basic region peptide. N and C indicate the N-terminal and C-terminal of the peptide, respectively. Half-circles at the C-terminal and filled ovals at the N-terminal represent β-cyclodextrin and the adamantyl group, respectively. White arrows denote the CRE half-site.

method[20] (0.1 mmol scale). *N*α-Fmoc-protected amino acid (0.25 mmol), PyBOP (0.25 mmol), *N*-methylmorpholine (0.375 mmol), and *N*-hydroxybenzotriazole (0.25 mmol) were used for coupling the amino acids. *N*α-Fmoc-protected amino acid ODhbt esters were used for Ser and Thr. Completion of the coupling reaction was monitored by the Kaiser test. The N-terminal of the peptide was acetylated with acetic anhydride (2.5 mmol). The peptide was cleaved from the resin by using an ice-cold TMSBr/thioanisole/TFA cleavage mixture containing 1.35 mL of TMSBr, 1.2 mL of thioanisole, 0.6 mL of ethanedithiol, 0.2 mL of *m*-cresol, and 7.48 mL of TFA. The solvent was removed, and the residue was dissolved in 30 mL of water and washed with diethyl ether. The mixture was passed through a Sephadex G-10 column with a 5% acetic acid solution. The peptide was purified by reversed-phase HPLC (column 20 × 250 mm, Ultron VX-Peptide; eluent A, 0.1% TFA−water; eluent B, 0.1% TFA−water containing 50% CH_3CN; linear gradient of 50−75% B over 40 min; flow rate 6 mL/min). The amino acid sequences of G1, G2, and G3 are shown in Figure 1. 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.

Synthesis of GAd Peptides. G1Ad was synthesized by a reaction of G1 (10 mg, 3.3 nmol) with iodoacetamide (18 mg, 0.1 mmol) in an aqueous DMF solution (pH 9, water:DMF = 3:2) at 0 °C under nitrogen. After 20 min, the reaction was quenched by addition of acetic acid, and the mixture was passed through a Sephadex G-10 column. Successive purification with reversed-phase HPLC (column 20 × 250 mm, Ultron VX-Peptide; eluent A, 0.2% TFA−water; eluent B, 0.2% TFA−water containing 50% CH_3CN; linear gradient of 55-75% B over 40 min; flow rate 6 mL/min) yielded pure G1Ad (8 mg) (yield 74%). **G1Ad**: amino acid analysis expected, Asx (1), Thr (1), Ser (1), Glx (2), Pro (1), Ala (6), Leu (2), Lys (3), Arg (6); found, Asx (1.0), Thr (0.87), Ser (0.53), Glx (2.13), Pro (1.04), Ala (6.35), Leu (2.08), Lys (2.41), Arg (6.08), MS (electrospray, 50% acetonitrile, 0.05% formic acid) calculated for [M^+] 3026.7, found 3025.3. G2Ad and G3Ad were synthesized in a similar manner from G2 and G3, respectively. **G2Ad**: amino acid analysis expected, Asx (2), Thr (1), Ser (1), Glx (2), Ala (6), Leu (2), Lys (3), Arg (6); found, Asx (2.0), Thr (0.86), Ser (0.55), Glx (2.04), Ala (6.09), Leu (1.86), Lys (2.28), Arg (5.76); MS (electrospray, 50% acetonitrile, 0.05% formic acid) calculated for [M^+] 3044.7, found 3043.7. **G3Ad**: amino acid analysis expected, Asx (2), Thr (1), Ser (1), Glx (2), Pro (1), Ala (6), Leu (2), Lys (3), Arg (6); found, Asx (2.0), Thr (0.29), Ser (0.89), Glx (2.62), Pro (1.05), Ala (6.07), Leu (2.10), Lys (2.25), Arg (6.94); MS (electrospray, 50% acetonitrile, 0.05% formic acid) calculated for [M^+] 3141.8, found 3141.1. The amino acid sequences of G1Ad, G2Ad, and G3Ad are shown in Figure 1. Stock solutions containing GAd were prepared by dissolving the peptide in TE buffer (1 mM Tris HCl, 0.1 mM EDTA, pH 8.0). Peptide concentrations were determined by quantitative amino acid analyses.

Synthesis of GAdCd Peptides. G1AdCd peptide was synthesized by a reaction of G1Ad (13 mg, 4.6 mmol) with mono(6-deoxy-6-iodo-β-cyclodextrin) (124 mg, 100 mmol) in an aqueous solution (pH 9) at 0 °C under nitrogen for 5 h. The reaction was quenched by addition of acetic acid, and the mixture was passed through a Sephadex G-10 column. G1AdCd was purified by reversed-phase HPLC (column 20 × 250 mm, Ultron VX-Peptide; eluent A, 0.05% TFA−water; eluent B, 0.05% TFA-water containing 50% CH_3CN; linear gradient of 30−65% B over 40 min; flow rate 6 mL/min). After HPLC purification, the mobile phase solution containing G1AdCd was neutralized by addition of triethylamine, and the resulting triethylammonium trifluoroacetate was removed by passing through a Sep-pak C18 cartridge (Waters). Lyophilization of the deionized solution containing G1AdCd yielded pure G1AdCd (5.5 mg) (yield 38%). **G1AdCd**: amino acid analysis expected, Asx (1), Thr (1), Ser (1), Glx (2), Pro (1), Ala (6), Leu (2), Lys (3), Arg (6); found, Asx (1.0), Thr (1.08), Ser (0.66), Glx (2.04), Pro (1.20), Ala (6.67), Leu (2.12), Lys (2.48), Arg (5.76); MS (electrospray, 50% acetonitrile, 0.05% formic acid) calculated for [M^+] 4086.7, found 4085.5. G2AdCd and G3AdCd were synthesized in a similar manner from G2 and G3, respectively. **G2AdCd**: amino acid analysis expected, Asx (2), Thr (1), Ser (1), Glx (2), Ala (6), Leu (2), Lys (3), Arg (6); found, Asx (2.0), Thr (0.86), Ser (0.55), Glx (2.04), Ala (6.09), Leu (1.86), Lys (2.28), Arg (5.76); MS (electrospray, 50% acetonitrile, 0.05% formic acid) calculated for [M^+] 4104.6, found 4104.5. **G3AdCd**: amino acid analysis expected, Asx (2), Thr (1), Ser (1), Glx (2), Pro (1), Ala (6), Leu (2), Lys (3), Arg (6); found, Asx (2.0), Thr (ND), Ser (1.24), Glx (1.89), Pro (0.97), Ala (6.23), Leu (1.97), Lys (2.49), Arg (ND); MS (electrospray, 50% acetonitrile, 0.05% formic acid) calculated for [M^+] 4201.8, found 4200.1. The amino acid sequences of G1AdCd, G2AdCd, and G3AdCd are shown in Figure 1. Stock solutions containing GAdCd were prepared by dissolving the peptide in TE buffer (1 mM Tris HCl, 0.1 mM EDTA, pH 8.0). Peptide concentrations were determined by quantitative amino acid analyses.

Synthesis and 5′-End Labeling of Oligonucleotides HS, T2, T2CE, and T3. Oligonucleotides were synthesized on a Milligen DNA synthesizer using a standard method and purified by reversed-phase HPLC on an Ultron VX-Nucleotide column (Sinwa Chemical Industry, 6 × 150 mm) with 0.1 M triethylammonium acetate−acetonitrile as the eluent. The oligonucleotides were labeled by kinase reaction using [γ-^{32}P]ATP (5000 Ci/mmol) and T4 polynucleotide kinase, followed by removal of unincorporated label over Sep-pak (Millipore Waters). The oligonucleotide was then denatured and annealed to a 4-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 HS (5′-CGGATGACACTGCTTTTTTTC-3′), T2 (5′-CGGATGACATGACTTTTTTTC-3′), T3 (5′-CGGATGACATGACATGACTTC-3′), and T2CE (5′-CGGATGACATTGCATGACTTC-3′). Sequences of only the single strand are shown.

Gel Mobility Shift Assay. Binding reactions were carried out in the presence of the indicated peptide 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, and 6% sucrose. 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/bisacrylamide), run in TBE buffer (20 mM Tris, 20 mM boric acid, and 0.1 mM EDTA) at 4 °C, and analyzed by autoradiography. The increase of the mobility-

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

shifted band was quantitated by the densitometry of the autoradiogram. The concentration of the peptide was determined by quantitative amino acid analysis with α-aminobutyric acid as the internal standard (Waters AccQ-Tag Chemistry Package, Millipore).

Measurement of CD Spectra. Spectra of the peptide 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, an appropriate concentration of peptide, 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 GCN4 Basic Region Peptides with Host and Guest Molecules. Both a DNA binding domain and a specific peptide−peptide interaction domain are required for a peptide that can bind DNA upon self-oligomerization. An oligopeptide derived from the basic region of the basic leucine zipper (bZIP) protein GCN4[21,22] was used as a DNA binding domain (Figure 1A), whereas β-cyclodextrin and the adamantyl group were used as a specific peptide−peptide interaction domain. In order to incorporate both guest and host molecules in the same peptide chain, an Fmoc-protected amino acid bearing an adamantyl group (Figure 1B) was synthesized. Condensation of the ϵ-amino group of *N*α-Boc-L-lysine with adamantaneacetic acid and successive removal of the Boc group and protection with the Fmoc group afforded *N*α-Fmoc-*N*ϵ-(adamantylacetyl)-L-lysine [Fmoc-Lys(Ad)]. The modified lysine residue bearing a guest molecule, Ad, at the ϵ-amino group was incorporated into the peptides at the positions indicated in Figure 1A.

Although the three-dimensional structures of a complex between GCN4 and its specific DNA sequences are available,[23−25] it is difficult to design a precise location of the peptide−peptide interaction domain. Thus, three peptides, G1, G2, and G3, were synthesized in which the positions of the adamantyl group were varied. β-Cd was attached at the C-terminal cysteine residue of the basic region peptides through the reaction of the SH group with 6-deoxy-6-iodo-β-cyclodextrin.[18,19] The resulting peptides G1AdCd, G2AdCd, and G3AdCd possess both host and guest molecules in the same peptide chain. Along with the GAdCd peptides, GAd peptides having no host molecule were synthesized as control peptides. The C-terminal cysteine residues of the GAd peptides were modified with iodoacetamide so that the GAd peptides lack an ability to form the host−guest inclusion complex.

Because the locations of the host and guest molecules determine the relative orientation of each monomer in the oligomer, all the GAdCd peptides will form tandemly repeated, head-to-tail homooligomers (Figure 1C). The parent basic region peptide recognizes a half-site of a palindromic CRE sequence (5′-ATGACGTCAT-3′) upon dimerization.[26,27] In addition, tandemly repeated basic region peptides have been shown to bind the direct repeat of the half-site DNA sequence when the peptides are covalently joined by the disulfide linkage.[11,12] Thus, the target DNA sequence for the oligomers of GAdCd would be a direct-repeat of the CRE half-site, 5′-$(ATGAC)_n$-3′.

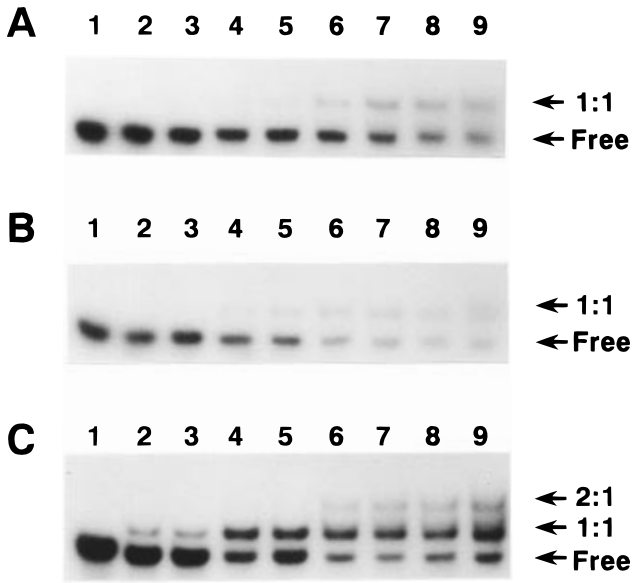


Figure 2. Autoradiograph showing titration of the gel shift[28,29] for G2Ad to HS (A), T2 (B), and T3 (C) oligonucleotides. Binding reactions were carried out as described in the Experimental Section. No G2Ad was added to the reaction in lane 1. G2Ad concentrations in lanes 2−9, respectively, were 1, 2, 4, 6, 10, 20, 30, and 50 nM.

DNA Binding of the GCN4 Basic Region Peptide Monomers. Binding of the GAd peptide to single (HS, 5′-ATGAC-3′), double (T2, 5′-ATGACATGAC-3′), and triple (T3, 5′-ATGACATGACATGAC-3′) direct-repeat sequence of the CRE half-site was first compared by titration of the gel shift.[28,29] G2Ad formed a 1:1 peptide−DNA complex with HS, T2, and T3 DNAs (Figure 2). The dissociation constant of G2Ad to HS DNA was 60 nM at 4 °C. A 2:1 G2Ad−DNA complex was observed with T3 DNA by increasing the concentration of G2Ad (Figure 2C). Such a 2:1 G2Ad−DNA complex was not observed for the T2 DNA under the conditions we have tested (Figure 2B). No obvious cooperativity was observed for the formation of the 2:1 G2Ad−T3 DNA complex. G1Ad and G3Ad peptides also form 1:1 complexes with HS, T2, and T3 DNAs. Dissociation constants of G1Ad and G3Ad to HS DNA were 40 and 90 nM, respectively, at the same conditions described for G2Ad. As was observed for G2Ad, the 2:1 peptide−DNA complexes were observed only with T3 DNA for both G1Ad and G3Ad.

It has been reported that a monomer of the basic region peptide does not bind to the target DNA sequences in a sequence-specific manner.[8] However, at the low temperature and the low salt conditions used here, even a monomer of the basic region peptide binds specifically to the CRE half-site. The specific complexes between GAd peptides and HS were not obtained with noncognate DNA sequences. Unlike the recently reported tethered complex of the monomeric basic region peptide and DNA,[30] the specific binding complex observed here is certainly unstable. In fact, the amount of total radio-labeled DNA fragment (free and distinct peptide complexes) reduced as the concentration of the basic region peptide was increased. This is possibly due to the formation of a nonspecific binding complex between the labeled DNA and the peptide(s). Such nonspecific complexes could dissociate during the gel electrophoresis, whereby reducing the amount of radio-labeled fragment at the free and the 1:1 complex bands.

(21) Hinnebush, A. G. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6442.
(22) Landschultz, W. H.; Johnson, P. F.; Adashi, E. Y.; Graves, B. J.; McKnight, S. L. *Genes Dev.* **1988**, *2*, 786.
(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.
(26) Sellers, J. W.; Vincent, A. C.; Struhl, K. *Mol. Cell. Biol.* **1990**, *10*, 5077.
(27) Kerppola, T. K.; Curran, T. *Curr. Opin. Struct. Biol.* **1991**, *1*, 71.
(28) Gamer, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047.
(29) Freid, M.; Crothers, D. M. *Nucleic Acids Res.* **1981**, *9*, 6505.
(30) Stanojevic, D.; Verdine, G. L. *Nat. Struct. Biol.* **1995**, *2*, 450.

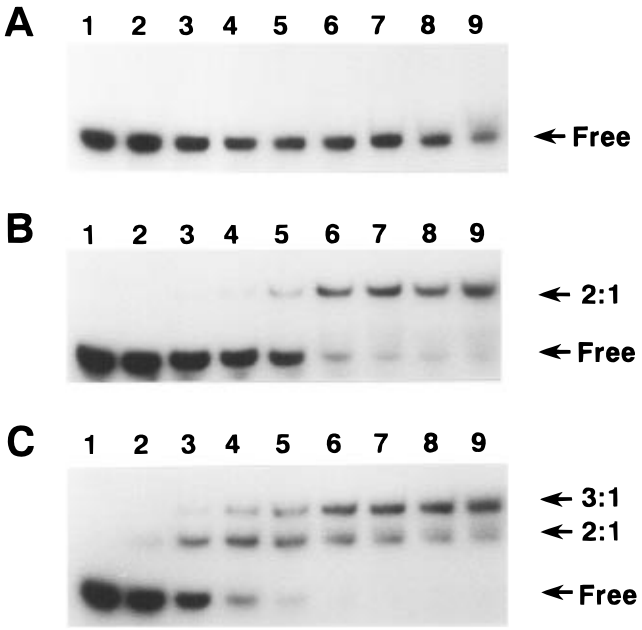


Figure 3. Autoradiograph showing the titration of the gel shift[28,29] for G2AdCd to HS (A), T2 (B), and T3 (C) oligonucleotides. Binding reactions were carried out as described in the Experimental Section. No G2AdCd was added to the reaction in lane 1. G2AdCd concentrations in lanes 2−9, respectively, were 1, 2, 4, 6, 10, 20, 30, and 50 nM.

DNA Binding of the β-Cyclodextrin-Modified GCN4 Basic Region Peptides. We next analyzed binding of G2AdCd to HS, T2, and T3 DNAs. In contrast to the case with G2Ad, G2AdCd did not form any binding complex with HS DNA within the concentration range tested for G2Ad (Figure 3A). The dissociation constant for the Ad−Cd complex is ca. 5 μM.[18] G2AdCd is expected to exist as a monomer in the absence of DNA. It is possible that two G2AdCd peptides would bind simultaneously to the T2 site, which possesses a tandem orientation of the 5′-ATGAC-3′ sequence, with positive cooperativity. With T2 DNA, only a mobility-shifted band corresponding to a 2:1 G2AdCd−DNA complex was observed, and neither a band corresponding to a 1:1 G2AdCd−DNA complex nor a multimer−DNA complex was detected (Figure 3B). The mobility-shifted band corresponding to a 2:1 G2AdCd−DNA complex shows almost the same mobility as that of the peptide dimer−DNA complex obtained in our previous study.[18,19] Because G2AdCd does not bind to HS, the formation of a 2:1 G2AdCd−DNA complex is sequence-specific.

DNA binding of G2AdCd to T3 DNA was examined next (Figure 3C). Again, no mobility-shifted band corresponding to the 1:1 G2AdCd−DNA complex was observed. However, in this case, two distinct mobility-shifted bands were observed. Judging from the mobility, the fast migrating band corresponds to the 2:1 complex of G2AdCd and T3 DNA. The slow band, which was not observed with T2 DNA, thus represents a 3:1 G2AdCd−T3 DNA complex. These results clearly show that the G2AdCd peptide binds DNA sequences comprising direct repeat of the CRE half-site with positive cooperativity. Moreover, DNA may function as a scaffold to promote the interaction between the bound peptides.

Effect of the Position of the Adamantyl Group on the Cooperative DNA Binding. A set of oligopeptides bearing both Ad and Cd groups in the same peptide chain was synthesized to analyze the effect of the relative positions of the guest molecule (Ad) to the host molecule (Cd) on the cooperative DNA binding. These peptides, G1AdCd, G2AdCd, and G3AdCd, possess essentially the same amino acid residues that are responsible for the direct interactions with the native GCN4 binding sequences (Figure 1) and the same host and guest molecules. Thus, these peptides differ from each other only in the position of the interpeptide interaction domain.

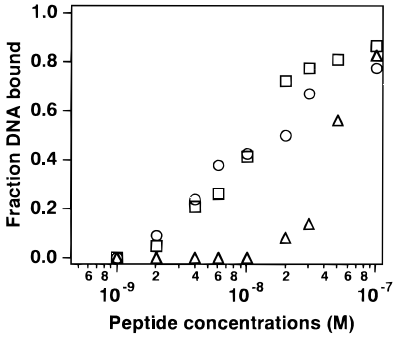


Figure 4. Semilogarithmic plots showing the fraction of ^{32}P-labeled T2 DNA bound to dimers of G1AdCd (open triangles), G2AdCd (open squares), and G3AdCd (open circles) as a function of total peptide concentration.

DNA binding of G1AdCd and G3AdCd was compared with that of G2AdCd. When HS DNA was used as a target, G1AdCd did not form any distinct peptide−DNA complexes within the concentrations tested (up to 100 nM). This result is consistent with that of G2AdCd. In contrast, G3AdCd formed a binding complex that migrates as a dimer−DNA complex. It is likely that one G3AdCd fit to the ATGAC site and the other G3AdCd molecule binds nonspecifically (half-specific and half-nonspecific) to either the N-terminal or C-terminal side of the specific complex.

All the GAdCd peptides formed the dimer−DNA complex with T2 DNA, although efficiencies for the formation of dimer−T2 DNA complexes differ from each other within the three peptides. Binding of G1AdCd, G2AdCd, and G3AdCd to T2 DNA was shown as plots illustrating the fraction of labeled T2 DNA bound to the GAdCd dimer as a function of the total peptide concentration (Figure 4). G3AdCd forms the 2:1 complexes as efficiently as G2AdCd does, while formation of the 2:1 complex between G1AdCd and T2 requires higher peptide concentrations. This is interesting because the G1Ad monomer shows the highest affinity with the half-site among the three GAd peptides. Thus, a subtle variation in the position of the adamantyl group affects the cooperative formation of the dimer−T2 DNA complex of GAdCd peptides. The geometry of the peptide−peptide interaction domain, not the stability of the DNA−peptide monomer interaction, plays a key role in the cooperative DNA binding of these peptides.

The position of the adamantyl group more significantly affects the binding of GAdCd peptides to T3 DNA. Both G1AdCd and G3AdCd afforded mobility-shifted bands corresponding to the 2:1 peptide−DNA complex with T3 DNA. The dimer complex of G3AdCd was formed more efficiently than that of G1AdCd (Figure 5). However, no distinct mobility-shifted band corresponding to the 3:1 peptide−DNA complex was observed for G1AdCd and G3AdCd. Among the six peptides G1Ad, G2Ad, G3Ad, G1AdCd, G2AdCd, and G3AdCd, only G2AdCd formed the 3:1 peptide−T3 DNA complex within the concentration range we have tested. The stabilities of the dimer−DNA and the trimer DNA complexes are highly influenced by geometrical constraints caused by the host−guest inclusion complex formation.

Sequence Selectivity of the Trimer−DNA Complex. The results described above clearly demonstrate that the geometry of the peptide−peptide interaction domain is an important factor

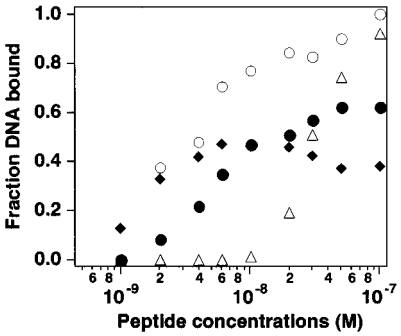


Figure 5. Semilogarithmic plots showing the fraction of ^{32}P-labeled T3 DNA bound to dimers of G1AdCd (open triangles), G2AdCd (filled squares), and G3AdCd (open circles) and the fraction of ^{32}P-labeled T3 DNA bound to a trimer of G2AdCd (filled circles) as a function of total peptide concentration.

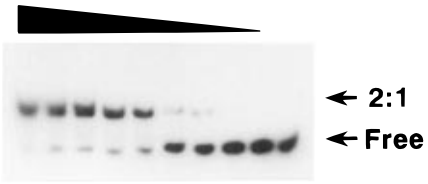


Figure 6. Autoradiograph showing the titration of the gel shift[28,29] for G2AdCd to T2CE oligonucleotides. Binding reactions were carried out as described in the Experimental Section. G2AdCd concentrations in lanes from the right were 0, 1, 2, 4, 6, 10, 20, 30, and 50 nM.

for the formation of peptide multimer−DNA complexes. In addition, the DNA sequence would function as a scaffold to promote the interaction between the bound peptides. As evident from the result of HS DNA and G3AdCd, half-specific and half-non-specific complexes could form depending on the nature of the peptide and DNA sequences. Sequence-selective formation of a trimer−DNA complex of G2AdCd was studied by using a mutant of T3 DNA (T2CE), in which the central half-site was changed into the 5′-ATTGC-3′ sequence. An autoradiogram showing the titration of G2AdCd binding to the T2CE sequence is shown in Figure 6. A mobility-shifted band indicative of formation of a 2:1 G2AdCd−DNA complex was observed as the concentration of the peptide was increased. However, the band corresponding to the 3:1 complex, which is clearly shown in the case of T3 DNA (Figure 3C), was not observed. Thus, the trimer−DNA complex formation of G2AdCd is quite sequence-specific. The 2:1 complex observed between G2AdCd and T2CE would contain at least two kinds of the half-specific and half-non-specific complexes. Because G2AdCd did not form a 2:1 complex with the HS DNA, it is likely that one molecule of G2AdCd fits nonspecifically to the 5′-ATTGC-3′ sequence and another molecule of G2AdCd binds specifically to one of the adjacent 5′-ATGAC-3′ sequences in the 2:1 complex between G2AdCd and the T2CE DNA.

Structures of GAd and GAdCd Peptides in the Absence or Presence of DNA. The parent GCN4 basic region peptide undergoes structural transition to a helical conformation upon binding to a specific DNA sequence, such as the palindromic 5′-ATGACGTCAT-3′ (CRE) sequence.[23−25,31] Thus, observed α-helicities of G2Ad and G2AdCd peptides in the presence of DNA will provide a good estimation for the specificity of the peptide−DNA interaction. The difference CD spectra of G2AdCd were compared in the presence of HS, T2, and T3 DNAs (Figure 7B). Interestingly, the intensity of the helical band for G2AdCd increased with increasing number of CRE

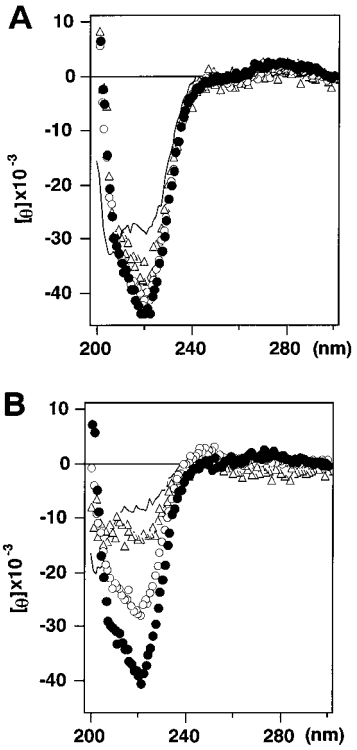


Figure 7. Circular dichroism difference spectra for G2Ad (A) and G2AdCd (B) in the absence or presence of various DNAs indicating that G2AdCd is helical when bound to the T2 or T3 DNA: spectra in the absence of DNA (solid line) and in the presence of HS (open triangles), T2 (open circles), and T3 (filled circles). Spectra of the peptides 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 as described in the Experimental Section.

half-site repeat sequences. That is, the intensity of the helical band increases in the order T3 > T2 > HS. In contrast, the difference CD spectra of G2Ad indicate that G2Ad is almost helical in the presence of a stoichiometric amount of the CRE half-site regardless of the number of repeat sequences (Figure 7A). Because the binding mixture contains a stoichiometric amount of the CRE half-site to the peptide, two molecules of G2AdCd would bind the doubly repeated half-site for T2 and three molecules for T3. The result indicates that the of G2AdCd becomes more helical with increasing number of half-site repeat sequences in the target DNA sequence.

Such a tendency was also observed for G1AdCd and G3AdCd (data not shown). All the GAdCd peptides showed highest helical contents in the presence of T3 DNA. However, only G2AdCd afforded the distinct trimer−DNA complex in the gel shift experiments. Because the CD spectral experiments were performed at micromolar concentrations of both peptides and DNA, even a noncooperative binding can be monitored. All the GAdCd peptides would bind the triple CRE half-sites in the T3 DNA in the helical form at such concentrations. Judging from the results obtained from the gel shift titration, three molecules of G2AdCd bind cooperatively to the T3 site. However, it is uncertain whether three bound molecules of G1AdCd or G3AdCd at the T3 site interact with each other through the host−guest complex formation.

Intramolecular Host−Guest Complex Formation of the GAdCd Peptides. In the absence of DNA, the intensity of the helical band of G2AdCd is weaker than that of G2Ad (Figure 7) even though the amino acid sequences for G2Ad and G2AdCd are exactly the same. Because the CD spectra of free G2AdCd remain unchanged within the concentration range from

(31) O'Neil, K. T.; Shuman, J. D.; Ampe, C.; DeGrado, W. F. *Biochemistry* **1991**, *30*, 9030.

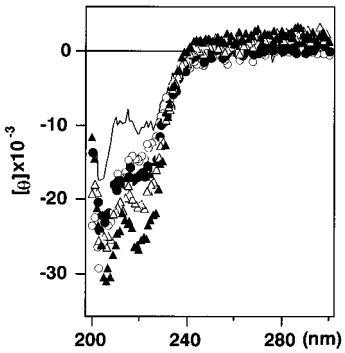


Figure 8. Circular dichroism spectra for G2AdCd in the absence or presence of 1-adamantaneacetic acid: spectra of G2AdCd in the absence of 1-adamantaneacetic acid (solid line) and in the presence of 10 μM (open cirles), 100 μM (filled circles), and 500 μM (open triangles) 1-adamantaneacetic acid and the spectrum of G2Ad (filled triangles). CD spectra were obtained as described in the Experimental Section.

0.5 to 10 μM, it is possible that the N-terminal adamantyl group of G2AdCd forms an intramolecular inclusion complex with the C-terminal Cd group without DNA.

In order to prove this, CD spectral titration of G2AdCd was performed by adding increasing amounts of 1-adamantaneacetic acid. The intramolecular inclusion complex between the adamantyl and the Cd groups of the peptide will be collapsed upon addition of an excess amount of 1-adamantaneacetic acid. As expected, the helicity of G2AdCd increases almost to that of free G2Ad by increasing the concentrations of 1-adamantaneacetic acid (Figure 8). Because the mean residue elipticity of free G2AdCd did not show any concentration dependency, it is quite likely that G2AdCd exists as a circular form in the absence of DNA. As illustrated in Figure 9, formation of such an intramolecular inclusion complex prevents transition of the monomeric G2AdCd peptide into a helical conformation, hence reducing the affinity with HS DNA as compared to G2Ad. In contrast, the helical content of G3AdCd is closer to that of G3Ad as compared to the cases of G1AdCd and G2AdCd (data not shown). It is possible that the intramolecular inclusion complex of G3AdCd is less stable than those of G1AdCd and G2AdCd. This explains why only G3AdCd affords the peptide dimer−DNA complex with the HS DNA.

Conclusions. The synthetic DNA binding peptides with self-oligomerization ability bind with positive cooperativity to the DNA sequences with multiple direct-repeat sequences. Introduction of an explicit interaction between the peptides, the host−guest interaction in the present study, into the DNA binding

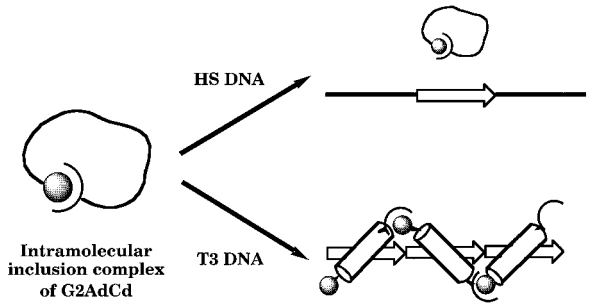


Figure 9. A schematic representation illustrating that G2AdCd forms the intramolecular inclusion complex in the absence of a specific DNA sequence and forms an α-helical trimer−DNA complex with T3 DNA. The intramolecular inclusion complex remains stable in the presence of HS DNA. Half-circles and filled ovals represent β-cyclodextrin and the adamantyl group, respectively. White arrows denote the CRE half-site. The cylinders represent the helical form of G2AdCd.

ability results in increased selectivity to the target DNA sequence. This is evident from the reduced affinity of GAdCd peptides to the HS DNA as compared to that of the GAd peptides. In our system, the stability of the cyclic peptide formed via an intramolecular inclusion complex is another key factor in reducing the affinity of GAdCd peptides to an isolated half-site as illustrated in Figure 9. The balance of intramolecular versus intermolecular interactions accounts for the binding selectivity.

The cooperative formation of dimer− and trimer−DNA complexes of the peptides is highly influenced by the geometry of the peptide−peptide interaction domain against the peptide−DNA complex. Although the DNA sequence would function as a scaffold to promote the interaction between the bound peptides, the geometrical constraints of the host−guest inclusion complex mainly determine the efficiencies for the formation of peptide multimer−DNA complexes. Thus, the observed high-selectivity was accomplished (i) by the cooperative nature of DNA binding and (ii) by reducing the stability of the non-specific DNA binding complex. Such strategies could be quite useful in designing novel sequence-specific DNA binding peptides.

Acknowledgment. We thank Dr. Hiroshi Ide and the reviewers for helpful comments. 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 and Culture, Japan.

JA953741M