J. Am. Chem. Soc. 1994, 116, 11137–11138

Communications to the Editor

Recognition of Nonpalindromic DNA Sequence by a Peptide Heterodimer with Artificial Dimerization Module

Masaru Ueno, Mamoru Sawada, Keisuke Makino, and Takashi Morii*

Department of Polymer Science & Engineering
Kyoto Institute of Technology
Matsugasaki, Sakyo-ku, Kyoto 606, Japan

Received July 6, 1994

Sequence-specific DNA binding of the gene regulatory protein is often mediated by dimeric species. Recent studies on the action of basic-helix-loop-helix and basic-leucine-zipper (bZIP) proteins have demonstrated an important role of the heterodimers in the control of specific gene activation.[1] While the mechanisms by which protein homodimers recognize the palindromic DNA sequences have emanated from the X-ray crystallographic studies of many protein–DNA complexes,[2,3] it has yet to be established whether the protein heterodimers could recognize nonpalindromic DNA sequences.[4] The heterodimer formation is a consequence of the specific protein–protein interaction between two different monomers. Thus, a prerequisite for DNA binding of the heterodimer is a module that appropriately controls such specific protein–protein interactions. We report here a sequence-specific DNA binding of peptide heterodimers with an artificial dimerization module[5] consisting of a β-cyclodextrin (β-Cd) and its guest compound. Because it is the host–guest inclusion complex that regulates the dimerization, attaching β-Cd to one peptide and a guest molecule to another peptide will allow a specific formation of peptide heterodimer.

We used peptides corresponding to the basic region of the yeast transcriptional activator GCN4[6] (G23) and an enhancer binding protein C/EBP[7] (C23) to study the DNA binding of heterodimer, since the basic region peptides of bZIP proteins alone are sufficient for the sequence-specific DNA binding when covalently[8–11] or noncovalently dimerized.[5] Modification of the C-terminal cysteines of the peptides with mono-6-deoxy-6-iodo-β-cyclodextrin (Cd) or *N*-bromoacetyl-1-adamantanemethylamine (Ad) afforded four different peptides (G23Ad, G23Cd, C23Ad, and C23Cd) that are capable of forming specific homo- (G23Ad/G23Cd and C23Ad/C23Cd) and heterodimers (C23Ad/G23Cd and C23Cd/G23Ad).[5,12] GCN4 and C/EBP are known to recognize palindromic sequences with a half-site of 5′-ATGAC-3′ (CRE) and 5′-ATTGC-3′ (CE), respectively.[13] Combination of these half-sites gives a nonpalindromic sequence (CE/CR), 5′-ATGACGCAAT-3′, for the target of the peptide heterodimer.

Specific recognition of the nonpalindromic DNA sequence by heterodimers was tested by using deoxyribonuclease (DNase) I footprinting[14] (Figure 1). Both C23Ad/G23Cd (lanes 4–6) and C23Cd/G23Ad (lanes 7–9) show specific protection at the CE/CR sequence. In contrast, G23Ad/G23Cd partially protects the canonical ATGAC half-site of the CE/CR sequence (lanes 11 and 12), and no protection is observed at the 3′-side of the

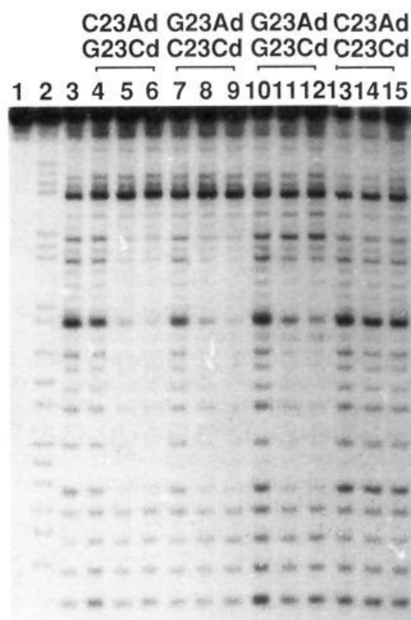

Figure 1. DNase I footprinting[14] pattern of the nonpalindromic CE/CR sequence by the homo- and heterodimers, indicating specific binding of the heterodimers C23Ad/G23Cd and C23Cd/G23Ad at the CE/CR sequence. The bracket denotes the nonpalindromic CE/CR sequence. Lane 1, no peptide; lane 2, Maxam–Gilbert chemical G + A reaction; lanes 3, DNase I digestion with no peptide; lanes 4–6, C23Ad/G23Cd (50, 150, and 300 nM); lanes 7–9, C23Cd/G23Ad (50, 150, and 300 nM); lanes 10–12, G23Ad/G23Cd (50, 150, and 300 nM); and lanes 13–15, C23Ad/C23Cd (50, 150, and 300 nM). DNA was 5′-end-labeled with polynucleotide kinase.[15] 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), 20 000 cpm singly 5′-^{32}P-end-labeled 46-mer DNA (5′-TCGAATTCCACAGTGAGAAATGACGCAATCCAGACTTA-GGATCCGC-3′, double-stranded) containing CE/CR sequence, and 1:1 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.

* To whom correspondence should be addressed. Present address: Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan.

(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) Vinson, C. R.; Hai, T.; Boyd, S. M. *Genes Dev.* **1993**, *7*, 1047.
(5) Ueno, M.; Murakami, A.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1993**, *115*, 12575.
(6) Hinnebush, A. G. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6442.
(7) Landschultz, W. H.; Johnson, P. F.; Adashi, E. Y.; Graves, B. J.; McKnight, S. L. *Genes Dev.* **1988**, *2*, 786.
(8) (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.
(9) (a) Cuenoud, B.; Schepartz, A. *Science* **1993**, *259*, 510. (b) Cuenoud, B.; Schepartz, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1154.
(10) (a) Park, C.; Campbell, J. L.; Goddard, W. A., III. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9094; (b) **1993**, *90*, 4892.
(11) Morii, T.; Shimomura, M.; Morimoto, M.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 1150.
(12) The amino acid sequences of C23 and G23 are Ac-NEYRVRRERNNIAVRKSRDKAKQC-NH_2 and Ac-DPAALKRARNTEAARRSRARKLQC-NH_2, respectively. Peptides G23 and C23 were synthesized as described previously (refs 5, 11). Structures of C23, C23Ad, and C23Cd were characterized by ^{1}H NMR spectroscopy and amino acid analysis. Peptide concentrations were determined by quantitative amino acid analysis with α-aminobutyric acid as an internal standard. Abbreviations for the amino acids are A, Ala; C, Cys; D, Asp; E, Glu; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; Y, Tyr.
(13) Kerppola, T. K.; Curran, T. *Curr. Opin. Struct. Biol.* **1991**, *1*, 71.
(14) Gallas, D. J.; Schmitz, A. *Nucleic Acids Res.* **1978**, *5*, 3157.
(15) Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1987.

0002-7863/94/1516-11137$04.50/0 © 1994 American Chemical Society

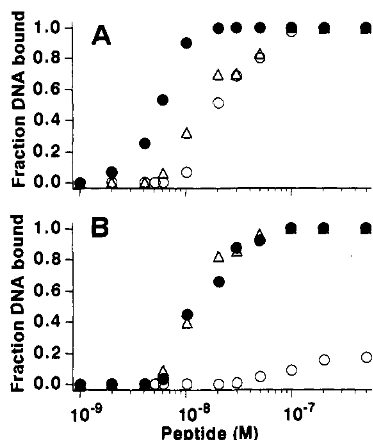

Figure 2. Titration of gel shift[17] for the basic region heterodimer C23Cd/G23Ad (A) and the homodimer C23Ad/C23Cd (B) to CECR21 (●), CRE21 (○), and CE21 (△) oligonucleotides. Extent of saturation for double-stranded oligonucleotides CECR21 (5'-CGG-ATTGCGTCAT)TTTTTTTC-3'), CRE21 (5'-CGGATGACGTCAT-TTTTTTTC-3'), and CE21 (5'-CGGATTGCGCAATTTTTTTTC-3') was obtained by averaging data from duplicate experiments. Binding reactions were carried out in the presence of the indicated peptides with <100 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, and 6% sucrose. 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/bisacrylamide) and run in TBE buffer (20 mM Tris, 20 mM boric acid, and 0.4 mM EDTA) at 4 °C, and the increase of the mobility-shifted band was quantitated by the densitometry of the autoradiogram.

nonmatched ATTGC sequence.[16] Another homodimer, C23Ad/C23Cd, shows no obvious protection against the DNase I digestion at the CE/CR sequence (lanes 13−15).

Binding selectivity of the homo- and heterodimers to the palindromic and nonpalindromic DNA sequences was compared by titration of the gel shift[17] (Figure 2). The heterodimer C23Cd/G23Ad preferentially binds to the CE/CR sequence over the palindromic CRE and CE sequences (Figure 2A). Homodimers G23Ad/G23Cd and C23Ad/C23Cd bind to the palindromic CRE[5] and CE sequences, respectively. However, both homodimers bind to the nonpalindromic CE/CR sequence as well, and, especially, C23Ad/C23Cd shows comparable affinity to the CE/CR and CE sequences (Figure 2B).[18] These results, together with the partial footprinting obtained for the G23 homodimer at CE/CR sequence, indicate an existence of half-matched binding complexes for DNA binding of the homo- and heterodimers.

Peptide structures of the matched and half-matched binding complexes were next analyzed by circular dichroism (CD) spectroscopy. Homodimers G23Ad/G23Cd[5] and C23Ad/C23Cd (Figure 3B) bind to each palindromic binding site, CRE and CE, respectively, predominantly in the α-helical conformation as observed for the native bZIP proteins.[19,20] With C23Cd/G23Ad, the intensity of the CD signal at 222 nm attributable to the helical band increased significantly in the presence of CECR21 (Figure 3A). The intensity of the helical band for C23Cd/G23Ad in the presence of CECR21 is comparable to that observed for C23Cd/C23Ad in the presence of CE21 (Figure 3B), suggesting that each monomer of C23Cd/G23Ad binds the CE/CR sequence in the α-helical conformation. The helical band intensity of the heterodimer in the presence of the palindromic sequence (CRE21 or CE21) is about half that observed for the CE/CR sequence (Figure 3A). This is also the case with the C23Ad/C23Cd in the presence of CECR21 (Figure 3B). Thus, the structure of the half-matched binding complex is partially helical, and most likely only the monomer at the matched half-site is in the α-helical conformation.

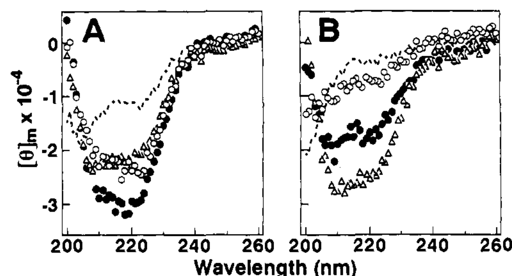

Figure 3. Circular dichroism difference spectra for the heterodimer C23Cd/G23Ad (A) and the homodimer C23Ad/C23Cd (B). The spectra indicate that the C23Cd/G23Ad complex is helical when bound to the nonpalindromic CE/CR sequence. Spectra in the absence of DNA (- - -) and in the presence of CECR21 (●), CRE21 (○), and CE21 (△) at 4 °C. Spectra of the dimers in the presence of oligonucleotides were calculated as the difference between the bound spectrum and a spectrum of the respective free oligonucleotide.[5,8,9] Samples contained 20 mM Tris-HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 4 μM each of Ad and Cd peptides, and 5 μM oligonucleotide duplex when present. $[\theta]_m$, the mean residue elipticity (deg cm^2 dmol^{-1}). Spectra were the average of 32 scans and were corrected with a spectrum of buffer alone but not smoothed.

In summary, our results demonstrate that the heterodimer of basic region peptides specifically recognizes the nonpalindromic DNA sequence in the helical structure. Efficiency of the sequence discrimination by the peptide dimers depends on a stability of the half-matched binding complex. However, structures of the matched and half-matched binding complexes differ from one another. Such subtle differences in structures would be critical for the subsequent specific protein−protein interactions that control the gene activation events.

Acknowledgment. This work was supported by Sasakawa Scientific Research Grant from the Japan Science Society and a 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. We thank Dr. Yukio Sugiura for reading the manuscript.

(16) In the gel shift experiment, mobility of the shifted band for CECR21 with G23Ad/G23Cd indicates that the binding complex is a 2:1 peptide−DNA complex. Thus, it is the G23Ad/G23Cd dimer, not a monomeric G23Ad or G23Cd, that binds at the CE/CR site under the conditions used in the footprinting experiments.

(17) (a) Garner, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047. (b) Freid, M.; Crothers, D. M. *Ibid* **1981**, *9*, 6505. Since the dissociation constant for the Ad/Cd dimerization module is ca. 5 μM (ref 5), it is possible that the loss of DNA binding of the peptide dimers at nanomolar concentration is due to loss of dimerization, not loss of binding of dimer. Stability of the dimer might be a key determinant for the DNA binding of the noncovalent dimers below nanomolar concentrations.

(18) Binding of bZIP protein homodimers to a nonpalindromic DNA site and a heterodimer to a palindromic DNA sequence has been reported recently (ref 4).

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

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