Arranging Quaternary Structure of Peptides by Cyclodextrin–Guest Inclusion Complex: Sequence-Specific DNA Binding by a Peptide Dimer with Artificial Dimerization Module

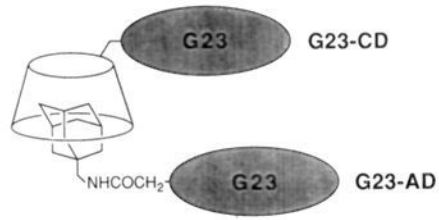
Masaru Ueno, Akira Murakami, Keisuke Makino, and Takashi Morii*

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

Received August 16, 1993

Arranging quaternary structures by dimerization of monomers is necessary for many sequence-specific DNA-binding proteins to become functional. Upon dimer formation, the DNA contacting regions of each monomer are positioned to a proper orientation that facilitates efficient sequence-specific recognition of DNA.[1] Several laboratories have reported model studies that demonstrate the importance of such steric constraints by using covalently bonded dimeric peptides.[2] However, natural proteins dimerize with noncovalent interactions, and the equilibrium governing the formation of dimers would be important for regulation.[3] Recent findings on the transcriptional control mechanisms by heterodimer formation for the basic leucine-zipper protein and the basic helix–loop–helix protein families also indicate that the specificity in dimer formation regulates the sequence-specific DNA binding of such proteins.[4] We report here synthetic oligopeptides that bind to specific DNA sequences upon noncovalent dimer formation. A peptide modified with β-cyclodextrin (β-CD) and a peptide with an adamantyl group form a heterodimer mediated by formation of an inclusion complex between the β-CD and the adamantyl group, and the heterodimer binds to a specific DNA sequence.

A 23-residue peptide (G23) derived from the DNA contacting region of the transcriptional activator protein GCN4,[5a] one of the large family of DNA-binding proteins characterized by a basic leucine-zipper structural motif, was used as a sequence-specific DNA-binding domain. GCN4 binds to the specific DNA sequence upon formation of dimer, which is mediated by a parallel-coiled coil assembly of each leucine-zipper domain.[5] Our design for effecting noncovalent dimer formation is a host–guest inclusion complex of β-CD and its guest compound. Modification of the C-terminal cysteine residue of G23 with mono-6-deoxy-6-iodo-β-cyclodextrin[6] afforded G23-CD, while modification with *N*-(bromoacetyl)-1-adamantanemethylamine gave G23-AD (Figure 1).[7] It has been shown that 1-adamantaneacetic acid forms a 1:1 inclusion complex in water with β-cyclodextrin with a dissociation constant of 5×10^{-5} M.[8] We have tested whether this level of interaction could effect the DNA binding of peptides containing the basic region sequence of GCN4.



G23 : Ac-DPAALKRARNTEAARRSRARKLQC-NH_2

Figure 1. Schematic representation of an inclusion complex of G23-AD and G23-CD and the amino acid sequence of G23.

Sequence-specific DNA binding of G23-CD and G23-AD are studied by gel mobility shift assay (Figure 2).[9] The basic region peptide with the adamantyl group, G23-AD, did not show any detectable binding to the ^{32}P-end-labeled DNA fragment containing native GCN4 binding sequence (CRE)[10] at the concentrations we have studied. In contrast, binding mixtures containing a 1:1 mixture of G23-AD and G23-CD show gradual appearance of a complex with lower electrophoretic mobility than CRE (lanes 8, 9, and 10, Figure 2A). The formation of a mobility-shifted band by the 1:1 mixture of G23-AD/G23-CD was inhibited on addition of free β-CD (lanes 14–16) or 1-adamantanemethylamine (lanes 11–13) to the binding mixture. These results strongly indicate that G23-AD and G23-CD form a heterodimer mediated by formation of an inclusion complex between β-CD and the adamantyl group.[11] Interestingly, G23-CD alone effected appearance of a mobility-shifted band with lower electrophoretic mobility than that of G23-AD/G23-CD (lanes 6 and 7), possibly through binding of G23-CD homodimer to CRE.[12] However, formation of the binding complex was far less efficient than that observed for the G23-AD/G23-CD mixture (lanes 7 and 10). Specific binding of the G23-AD/G23-CD complex to CRE was confirmed by gel mobility shift competition assay[9] with nonradiolabeled CRE, an oligonucleotide containing an AP-1 site,[13] and an oligonucleotide without native GCN4 binding sequence[14] (Figure S2).

Binding of the G23-AD/G23-CD mixture to the CRE sequence was further characterized by deoxyribonuclease I (DNase I)

* To whom correspondence should be addressed.

(1) (a) Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1984**, *53*, 293. (b) Harrison, S. C.; Aggarwal, A. K. *Ibid.* **1990**, *59*, 933. (c) Pabo, C. O.; Sauer, R. T. *Ibid.* **1992**, *61*, 1053.

(2) (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. (c) Morii, T.; Simomura, M.; Morimoto, S.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 1150. (d) Cuenoud, B.; Schepartz, A. *Science* **1993**, *259*, 510. (e) Cuenoud, B.; Schepartz, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1154. (f) Park, C.; Campbell, J. L.; Goddard, W. A. III. *Ibid.* **1992**, *89*, 9094. (g) Park, C.; Campbell, J. L.; Goddard, W. A. III *Ibid.* **1993**, *90*, 4892.

(3) (a) Ptashne, M. *A Genetic Switch*; Blackwell Scientific Publications and Cell Press: Palo Alto, CA, 1986. (b) Jones, N. *Cell* **1990**, *61*, 9. (c) Brenowitz, M. B.; Jamison, E.; Majumdar, A.; Adhya, S. *Biochemistry* **1990**, *29*, 3374. (d) Kim, B.; Little, J. W. *Science* **1992**, *255*, 203. (e) Brown, B. M.; Sauer, R. T. *Biochemistry* **1993**, *32*, 1354.

(4) (a) Frankel, A.; Kim, P. S. *Cell* **1991**, *65*, 717. (b) O'Shea, E. K.; Rutkowski, R.; Kim, P. S. *Cell* **1992**, *68*, 699.

(5) (a) Vinson, C. R.; Sigler, P. B.; McKnight, S. L. *Science* **1989**, *246*, 911. (b) Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223.

(6) Melton, L. D.; Slessor, K. N. *Carbohydr. Res.* **1971**, *18*, 29.

(7) Peptide G23 was synthesized as described previously (ref 2c). G23-CD was synthesized by a reaction of G23 with mono-6-deoxy-6-iodo-β-cyclodextrin (8-fold molar excess) in aqueous solution (pH 9) at 0 °C under nitrogen. The reaction was quenched on addition of acetic acid, and the mixture was passed through a Sephadex G-10 column. Successive purification with a reverse-phase HPLC yielded pure G23-CD. G23-AD was obtained in a similar manner, except the reaction was carried out with *N*-(bromoacetyl)-1-adamantanemethylamine (15-fold molar excess) in aqueous DMF solution (pH 9, water:DMF = 3:2). Structures of G23, G23-AD, and G23-CD were characterized by ^{1}H NMR spectroscopy and amino acid analysis. 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.

(8) Eftink, M. R.; Andy, M. L.; Bystrom, K.; Perlmutter, H. D.; Kristol, D. S. *J. Am. Chem. Soc.* **1989**, *111*, 6765.

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

(10) A 21-bp double-stranded oligonucleotide, CRE, containing GCN4 target site (ATGACGTCAT) is 5′-CGG*ATGACGTCAT*TTTTTTTTC-3′ and its complementary strand.

(11) Half-maxmal binding of the 1:1 mixture of G23-AD/G23-CD to CRE occurs at each concentration of 30 nM as revealed by the gel mobility shift titration experiments. A more detailed description of these binding experiments will be presented elsewhere (M.U., A.M., K.M., and T.M., in preparation).

(12) This is expected since the molecular weight of the G23-CD homodimer is larger than that of the G23-AD/G23-CD heterodimer. The binding of G23-CD is competed away with free adamantaneacetic acid (Figure S1). No mobility-shifted band corresponding to the homodimer of G23-CD is observed in the binding mixture containing a 1:1 mixture of G23-AD/G23-CD.

(13) A 20-bp double-stranded oligonucleotide, GRE, containing GCN4 target site (ATGACTCAT) is 5′-CGG*ATGACTCAT*TTTTTTTTC-3′ and its complementary strand.

(14) A 25-bp double-stranded oligonucleotide, NON, is 5′-GATC-CCCCCAACACCTGCTGCCTGA-3′ and its complementary strand.

0002-7863/93/1515-12575$04.00/0

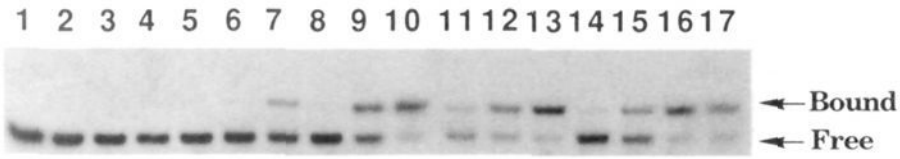


Figure 2. Gel mobility shift assay showing the binding of G23-AD/G23-CD to ^{32}P-end-labeled CRE, and inhibitions of the binding of G23-AD/G23-CD on addition of β-CD or adamantanemethylamine. Lane 1, no peptide; lanes 2, 3, and 4, 10, 50, and 100 nM G23-AD; lanes 5, 6, and 7, 10, 50, and 100 nM G23-CD; lanes 8, 9, and 10, 10, 50, and 100 nM each of G23-AD and G23-CD; lanes 11–17 contain 200 nM each of G23-AD and G23-CD; lanes 11–13, 1000, 100, and 10 μM adamantanemethylamine; lanes 14–16, 1000, 100, and 10 μM β-CD. Binding reactions were carried out in the presence of the indicated peptide with <1 nM labeled CRE in a binding mixture (20 μL) containing 20 mM Tris-HCl (pH 7.5), 4 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 20 μM poly(dI-dC), 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 8% nondenaturing acrylamide gel (29:1 acrylamide/bis-acrylamide), run in 1/5 TBE buffer (20 mM Tris, 20 mM boric acid, and 0.4 mM EDTA) at 4 °C, and analyzed by autoradiography.

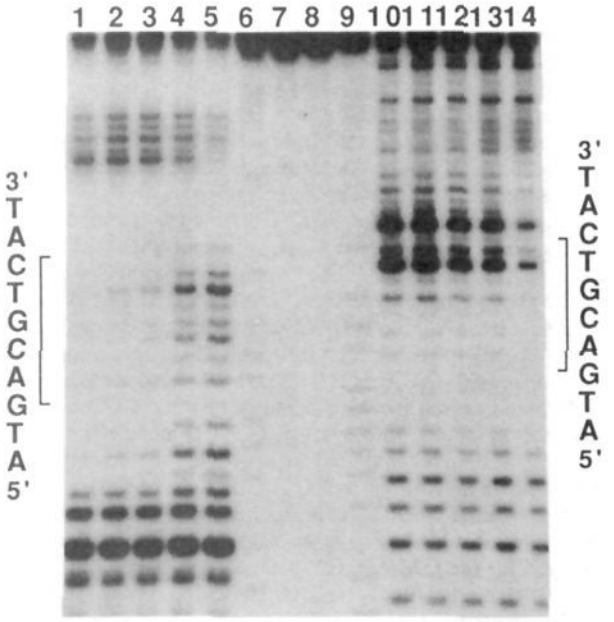


Figure 3. DNase I footprint of the CRE site by G23-AD/G23-CD complex. Brackets denote the native GCN4 binding sequence. Lanes 1–7, 5′-^{32}P-end-labeled 46-mer top strand; lanes 8–14, 5′-^{32}P-end-labeled 46-mer bottom strand; lanes 7 and 8, intact DNA; lanes 6 and 9, Maxam–Gilbert chemical G+A reaction; lanes 1–5 and 10–14, DNA digested with DNAseI; lanes 5 and 10, no peptide; lanes 4 and 11, 50 mM peptide; lanes 3 and 12, 100 nM peptide; lanes 2 and 13, 200 nM peptide; lanes 1 and 14, 300 nM peptide. 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 ^{32}P-end-labeled DNA, and a 1:1 mixture of G23-AD/G23-CD, where indicated, in 20 μL total volume. Nuclease digestion was initiated on addition of 1 unit of DNase I (Takara) 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, run on 12% sequencing acrylamide gel, and analyzed by autoradiography.

footprinting[15] and circular dichroism (CD) spectroscopy. As shown in Figure 3, the G23-AD/G23-CD mixture showed distinct protection at the whole native GCN4 binding sequence (5′-ATGACGTCAT-3′) for both strands. The basic region of GCN4 has disordered structure in the absence of a specific DNA sequence but is structured upon binding to a specific DNA sequence.[2a,b,16] The negative CD signal at 222 nm, corresponding to an α-helical structure, was increased significantly on addition of CRE to the 1:1 mixture of G23-AD/G23-CD (Figure 4). The difference

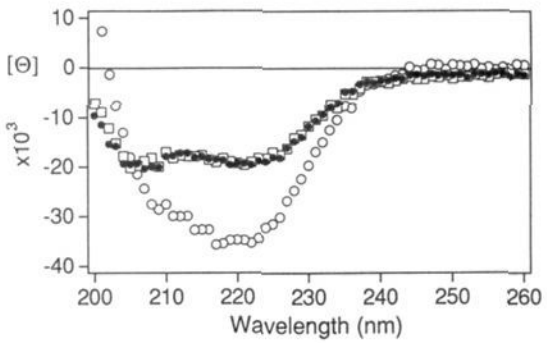


Figure 4. Circular dichroism spectroscopy results for the G23-AD/G23-CD complex bound to CRE. The spectra indicate that the G23-AD/G23-CD complex is helical when bound to CRE. G23-AD/G23-CD (●), G23-AD/G23-CD bound to CRE (○), and G23-AD/G23-CD in the presence of NON (□). Spectra of G23-AD/G23-CD in the presence of CRE or NON were calculated as the difference between the bound spectrum and free CRE or NON spectrum. CD spectra were obtained 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 each of G23-AD and G23-CD, and 5 μM CRE or NON when present.

spectrum indicates that G23-AD and G23-CD become almost helical upon binding to CRE, as has been observed for native GCN4[16] and the covalently bonded GCN4 basic region dimer.[2a] No such change of CD signal at 222 nm was observed on addition of the nonspecific oligonucleotide NON.[14] Thus the peptide G23 of G23-AD/G23-CD heterodimer contacts to DNA with the same recognition mechanisms as observed for the basic region of native GCN4.[5b,16]

In summary, we have shown that β-cyclodextrin and its guest compound can be used to generate an efficient heterodimerization module. Our strategy of coupling the host–guest inclusion complex with small peptides would be quite useful for arranging functional quaternary structures of artificial proteins in aqueous solutions.

Acknowledgment. This work was supported by Sasakawa Scientific Research Grant from the Japan Science Society to T.M. We thank Drs. Jacqueline Barton and Peter Kim for helpful discussions.

Supplementary Material Available: Characterization details for autoradiograms showing a competition of the binding of G23-CD to CRE with free adamantaneacetic acid (Figure S1) and a band shift competition assay showing specific binding of G23-AD/G23-CD to CRE (Figure S2) (5 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

(15) Gallas, D. J.; Schmitz, A. *Nucleic Acids Res.* **1978**, *5*, 3157.
(16) (a) Weiss, M. A.; Ellenberger, T.; Wobbe, C. R.; Lee, J. P.; Harrison, S. C.; Struhl, K. *Nature* **1990**, *347*, 575. (b) Patel, L.; Abate, C.; Curran, T. *Ibid* **1990**, *347*, 572.