J|A|C|S
COMMUNICATIONS

Published on Web 05/14/2002

Novel Strategy for the Design of a New Zinc Finger: Creation of a Zinc Finger for the AT-Rich Sequence by α-Helix Substitution

Makoto Nagaoka, Yoshihide Doi, Jun Kuwahara,† and Yukio Sugiura*

Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Received February 7, 2002

The C_2H_2-type zinc finger is one of the most typical DNA binding motifs found in various DNA binding proteins and has a $\beta\beta\alpha$ structure formed by the zinc complexation with two cysteines and two histidines as ligands.[1] The structural analyses of several C_2H_2-type zinc finger−DNA complexes revealed the characteristics of this motif in DNA binding as follows: (1) the zinc finger structure is repeated by the connection with a specific linker, (2) each finger structure binds to a three-base-pair subsite of DNA with its α-helix facing toward the major groove, (3) the amino acid residues at four key positions in the α-helix of each finger unit make a 1:1 contact with the DNA bases at specific positions, and (4) the overall arrangement of the peptide is antiparallel to the primary interacting strand of DNA.[2] This modular DNA recognition mode provides the expectation of the development of a DNA recognition code for the zinc finger. In practice, libraries of zinc finger variants have been created by designed and phage display strategies.[3] The attempt has had moderate success, but the sequences which are not recognized by any zinc fingers still remain.

Most of the natural proteins with C_2H_2-type zinc fingers bind to GC-rich sequences by mainly recognizing guanines. Human transcription factor Sp1 also contains three C_2H_2-type zinc fingers in the C-terminal domain (fingers 1−3) and binds to the GC box, 5′-GGG GCG GGGC-3′.[4] On the other hand, the C-terminal three of six C_2H_2-type zinc fingers (fingers 4−6) derived from *Drosophila* transcription factor CF2-II possess a sequence preference for the AT-rich element, 5′-GTA TAT ATA-3′.[5] In this communication, a novel zinc finger peptide binding to the AT-rich sequence was created by α-helix substitution between these proteins (Figure 1A). Figure 1B shows the primary sequences and target DNA subsites of the three zinc fingers of Sp1(zf123)[6] and Sp1HM. Sp1HM is a helix-substituted chimeric mutant between Sp1(zf123) and CF2-II and its α-helix of each finger is replaced by that of fingers 4−6 of CF2-II.[7] The folding property of Sp1HM was analyzed by the measurement of the circular dichroism spectrum (Figure 2A). The spectrum of Sp1HM is comparable with those of Sp1(zf123) and the single zinc finger peptide.[8] Negative Cotton effects with a minimum at 206 nm and a shoulder around 222 nm suggest that Sp1HM has an ordered secondary structure, namely α-helices.[9]

The DNA binding affinity and specificity of Sp1HM for GC: 5′-GGG GCG GGGC-3′ and AT: 5′-GTA TAT ATA-3′ were evaluated by gel mobility shift assays, and the dissociation constants for each substrate were determined (Figure 2B, Table 1).[10] It is expected that Sp1HM has a preference for AT since its α-helix for DNA recognition in each finger is derived from that of CF2-II. Sp1HM binds to AT with 3.2 and 1330 nM K_d values in the absence and presence of competitor DNA, respectively. In contrast, the K_d values of Sp1(zf123) for GC in the absence and presence of competitor DNA were 4.0 and 46.1 nM, respectively. No evident binding complex for Sp1(zf123)-AT or Sp1HM-GC was detected under our experimental conditions. These results suggest that Sp1HM binds to AT with a high affinity and specificity. The zinc finger peptide for the TATA box (5′-GCT ATA AAA-3′) was

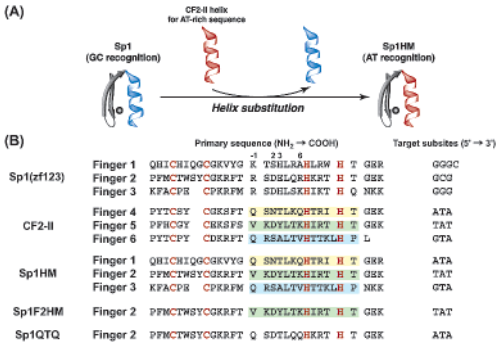

Figure 1. (A) Strategy of α-helix substitution for the design of a new zinc finger. (B) Primary sequences and target DNA triplets of zinc fingers of Sp1(zf123), CF2-II, Sp1HM, Sp1F2HM, and Sp1QTQ. Cysteines and histidines for zinc ligands are shown as red characters. Yellow, green, and blue colored boxes indicate the helical regions derived from the fingers 1, 2, and 3 of CF2-II, respectively. The sequences of fingers 1 and 3 of Sp1F2HM and Sp1QTQ are identical with those of fingers 1 and 3 of Sp1-(zf123), respectively.

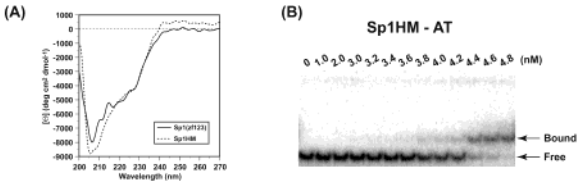

Figure 2. (A) Circular dichroism spectra of Sp1(zf123) (solid line) and Sp1HM (dotted line) were recorded on a Jasco J-720 spectropolarimeter in 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM dithiothreitol, and 5 μM peptide at 20 °C. (B) Gel mobility shift assay for determining the dissociation constant of Sp1HM in the absence of competitor DNA. The peptide concentrations are shown in the figure.

Table 1. Dissociation Constants (K_d) for Sp1(zf123) and Sp1HM in the Presence and Absence of Competitor DNAs

	K_d (nM)[a]			
binding site	Sp1(zf123)	Sp1(zf123) + CT DNA[b]	Sp1HM	Sp1HM + CT DNA
GC	4.0 ± 0.1	46.1 ± 2.9	NBD	NBD
AT	NBD[c]	NBD	3.2 ± 1.6	1330 ± 163

[a] Dissociation constants were estimated by titration using a gel mobility shift assay. Values are averages of three independent determinations with standard deviations. [b] CT DNA, calf thymus DNA. [c] NBD, no binding detected.

* To whom correspondence should be addressed. E-mail: sugiura@scl.kyoto-u.ac.jp.
† Present address: Faculty of Pharmaceutical Sciences, Tokushima University, Tokushima 770-8505, Japan.

10.1021/ja025856d CCC: $22.00

created by Greisman and Pabo.[12] They used the three-zinc-finger domain derived from Zif268 as a framework and succeeded in selecting the $TATA_{ZF}$ by the sequential selection strategy based on the phage display technique. The $TATA_{ZF}$ binds to the TATA box with a 0.12 nM dissociation constant,[12b] indicating that the DNA binding affinity of the zinc finger peptides for AT-rich sequence is estimated at less than nanomolar order.[13] These results imply that our rational design is also effective and convenient for creating zinc finger peptides with a high affinity comparable with that of zinc fingers selected by the phage display technique.

Recently, the crystal structure of a $TATA_{ZF}$−TATA box complex was resolved by Pabo et al.[14] In the complex, $TATA_{ZF}$ makes base contacts unpredictable from the conventional DNA recognition rule of zinc fingers. Namely, the amino acids at two helical positions, 1 and 5, are essential for binding affinity and specificity in addition to those at the four key positions. The importance of such amino acids in the DNA binding affinity and specificity has also been reported in the DNA binding of CF2-II,[5c] suggesting that Sp1HM has similar unusual base contacts.

Despite the increase in the amino acid−DNA interactions, the present results clearly demonstrate that the affinity of Sp1HM for its target DNA sequence is almost the same as that of Sp1(zf123). Moreover, $TATA_{ZF}$ binds to the target DNA less tightly than Zif268.[12] In comparison with the Zif268−DNA complex in which the main interaction force for base recognition is the hydrogen bond, a hydrophobic or van der Waals interaction as well as a hydrogen bond are dominant in the $TATA_{ZF}$−DNA complex. The binding energy of the van der Waals interaction is about 10-fold smaller than that of the hydrogen bond, indicating that the increase in the number of interactions between the amino acids and bases is not sufficient for the DNA binding of $TATA_{ZF}$ with high affinity. Along with the interaction force, the ascendancy in the bidentate guanine recognition by arginine, in which two hydrogen bonds are formed, may provide a higher DNA binding affinity of Zif268. In view of our peptide design, this consideration is also applicable to our system and accounts for the almost same or lower binding affinity of the zinc finger peptide for AT-rich sequence in comparison with that for the GC-rich sequence.

The creation of the zinc finger binding to the sequences alternating between the AT- and GC-rich subsites is one of the important subjects in zinc finger−DNA interaction. To examine the validity of our strategy, we prepared the zinc finger peptides for such sequences by helix substitution (Sp1F2HM) and the point mutations of four key positions (Sp1QTQ) and then examined the DNA binding of the peptides (Figure 1B).[15] Sp1F2HM was strongly bound to its target sequence ($K_d = 3.6 \pm 0.1$ nM), although Sp1QTQ formed no stable complexes with GC, AT, or the target sequence (data not shown). However, the second finger of Sp1F2HM could not distinguish 5′-TAT-3′ from 5′-ATA-3′. Of interest is the contrast between Sp1F2HM and Sp1QTQ. These results strongly indicate that our strategy of α-helix substitution is also effective in the creation of the zinc finger peptide binding to the sequences alternating between the AT- and GC-rich subsites, although more effort has to be made to produce those with a higher specificity.

In this communication, we propose a novel strategy for the creation of new zinc fingers by helix substitution. Based on this strategy, we have succeeded in producing not only the zinc finger peptide for the AT-rich sequence, but also the initial one for the sequence alternating between the AT- and GC-rich subsites. The strategy is evidently effective and is also more convenient than the phage display method. Consequently, our design method is widely applicable to creating zinc finger peptides with novel binding specificities.

Acknowledgment. This study was supported in part by Grants-in-Aid for COE Project "Element Science" (12CE2005) and Scientific Research (12470505, 12793008, and 13557210) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

References

(1) (a) Párraga, G.; Horvath, S. J.; Eisen, A.; Taylor, W. E.; Hood, L.; Young, E. T.; Klevit, R. E. *Science* **1988**, *241*, 1489−1492. (b) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635−637.

(2) (a) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809−817. (b) Pavletich, N. P.; Pabo, C. O. *Science* **1993**, *261*, 1701−1707. (c) Fairall, L.; Schwabe, J. W. R. *Nature* **1993**, *366*, 483−487. (d) Elrod-Erickson, M.; Rould, M. A.; Nekludova, L.; Pabo, C. O. *Structure* **1996**, *4*, 1171−1180. (e) Kim, C. A.; Berg, J. M. *Nat. Struct. Biol.* **1996**, *3*, 940−945. (f) Houbaviy, H. B.; Usheva, A.; Shenk, T.; Burley S. K. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13577−13582. (g) Wuttke, D. S.; Foster, M. P.; Case, D. A.; Gottesfeld, J. M.; Wright, P. E. *J. Mol. Biol.* **1997**, *273*, 183−206. (h) Nolte, R. T.; Conlin, R. M.; Harrison, S. C.; Broun, R. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2938−2943. (i) Elrod-Erickson, M.; Benson, T. E.; Pabo, C. O. *Structure* **1998**, *6*, 451−464.

(3) (a) Rebar, E. J.; Pabo, C. O. *Science* **1994**, *263*, 671−673. (b) Jamieson, A. C.; Kim, S.-H.; Wells, J. A. *Biochemistry* **1994**, *33*, 5689−5695. (c) Choo, Y.; Klug, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11163−11167. (d) Choo, Y.; Klug, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11168−11172. (e) Wu, H.; Yang, W.-P.; Barbas, C. F., III *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 344−348. (f) Jamieson, A. C.; Wang, H.; Kim, S.-H. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12834−12839.

(4) (a) Gidoni, D.; Dynan, W. S.; Tjian, R. *Nature* **1984**, *312*, 409−413. (b) Kadonaga, J. T.; Jones, K. A.; Tjian, R. *Trends Biochem.* **1986**, *11*, 20−23. (c) Kadonaga, J. T.; Carner, K. R.; Masiarz, F. R.; Tjian, R. *Cell* **1987**, *51*, 1079−1090.

(5) (a) Hsu, T.; Gogos, J. A.; Kirsh, S. A.; Kafatos, F. C. *Science* **1992**, *257*, 1946−1950. (b) Gogos, J. A.; Hsu, T.; Bolton, J.; Kafatos, F. C. *Science* **1992**, *257*, 1951−1955. (c) Gogos, J. A.; Jin, J.; Wan, H.; Kokkinidis, M.; Kafatos, F. C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2159−2164.

(6) (a) Nagaoka, M.; Sugiura, Y. *Biochemistry* **1996**, *35*, 8761−8768. (b) Uno, Y.; Matsushita, K.; Nagaoka, M.; Sugiura, Y. *Biochemistry* **2001**, *40*, 1787−1795.

(7) Sp1HM was cloned by using pEVSp1(530−623)[6a] as a template for the polymerase chain reaction and expressed and purified according to a previous procedure.[6b]

(8) (a) Kuwahara, J.; Coleman, J. E. *Biochemistry* **1990**, *29*, 8627−8631. (b) Frankel, A. D.; Berg, J. M.; Pabo, C. O. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4841−4845. (c) Hori, Y.; Suzuki, K.; Okuno, Y.; Nagaoka, M.; Futaki, S.; Sugiura, Y. *J. Am. Chem. Soc.* **2000**, *122*, 7648−7653.

(9) As for the molecular ellipticities at 206 nm, the values of Sp1(zf123) and Sp1HM were −7921 and −8770 nm, respectively. For a leading reference, see: *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996.

(10) Gel mobility shift assays were performed under the following conditions. Each reaction mixture contained 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.05% Nonidet P-40, 5% glycerol, the ^{32}P-end-labeled substrate DNA fragment (∼50 pM), and 0−2.5 μM zinc finger peptide. For the experiment in the presence of a competitor DNA, calf thymus DNA (25 ng/μL) was also added. After incubation at 20 °C for 30 min, the sample was run on a 12% polyacrylamide gel with Tris-borate buffer at 20 °C. The bands were visualized by autoradiography and quantified with ImageMaster 1D Elite software (version 3.01). The dissociation constants (K_d) of the Sp1 peptide−DNA complexes were estimated based on a previously reported procedure.[11]

(11) Yokono, M.; Saegusa, N.; Matsushita, K.; Sugiura, Y. *Biochemistry* **1998**, *37*, 6824−6832.

(12) (a) Greisman, H. A.; Pabo, C. O. *Science* **1997**, *275*, 657−661. (b) Wolfe, S. A.; Greisman, H. A.; Ramm, E. I.; Pabo, C. O. *J. Mol. Biol.* **1999**, *285*, 1917−1934.

(13) Although we examined the DNA binding of Sp1HM under the same buffer conditions as that reported by Greisman and Pabo,[12a] the high affinity binding of Sp1HM was not observed. This evidence demonstrates that the DNA binding affinity of Sp1HM is affected by the buffer conditions. Consequently, the affinity may increase much more under the optimized buffer conditions.

(14) Wolfe, S. A.; Grant, R. A.; Elrod-Erickson, M.; Pabo, C. O. *Structure* **2001**, *9*, 717−723.

(15) Sp1QTQ has glutamine, threonine, and glutamine at positions −1, 3, and 6 in the second finger, respectively, and is expected to bind to 5′-GGG ATA GGG-3′. On the other hand, Sp1F2HM is a mutant in which only the second finger of Sp1(zf123) is changed to that of Sp1HM and its target sequence is 5′-GGG TAT GGG-3′.

JA025856D