The plant zinc-finger protein ZPT2-2 has a unique mode of DNA interaction

Ken-ichi Yoshioka,¹ Setsuko Fukushima,¹ Toshimasa Yamazaki,² Michiteru Yoshida,¹ and Hiroshi Takatsuji³*

¹ Department of Biological Science and Technology, Science University of Tokyo, 2641 Yamazaki, Noda, Chiba 278-8510, Japan

Departments of ² Biotechnology and ³ Plant Physiology, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan

* To whom correspondence should be addressed.

Telephone number: (+81) 298-38-8383

Fax number: (+81) 298-38-8383

E-mail: takatsuh@nias.affrc.go.jp

Running title: Unique DNA interaction of a plant zinc-finger protein
Summary

ZPT2-2 is a DNA-binding protein of petunia that contains two canonical TFIIIA-type zinc-finger motifs separated by a long linker. We previously reported that ZPT2-2 bound to two separate AGT core sites, with each zinc finger making contact with each core site. Here we present our further characterization of ZPT2-2 by using SAAB imprinting and surface plasmon resonance (SPR) analyses; together, these assays revealed some unusual features of the interaction between ZPT2-2 and DNA. These experiments allowed us to conclude that 1) the optimal binding sequence for the N-terminal zinc finger is AGC(T), and that of the C-terminal one is CAGT; 2) multiple arrangements of the two core sites accommodate binding; and 3) the spacing between the two core sites affects the binding affinity. In light of these observations, we propose a new model for the DNA–ZPT2-2 interaction. Further, consistent with this model, a high-affinity binding site for ZPT2-2 was found in the promoter region of the ZPT2-2 gene. This site may serve as a cis-element for the autoregulation of ZPT2-2 gene expression.
Unique DNA-interaction of a plant zinc-finger protein

Introduction

The TFIIIA (Cys$_2$/His$_2$)-type zinc-finger proteins, first discovered in transcription factor IIIA of Xenopus (1), represent an important class of eukaryotic transcription factors. To date, numerous genes have been found that encode this type of zinc-finger motif, and many of their products have been implicated in various regulatory roles. The TFIIIA-type zinc finger tetrahedrally coordinates a zinc atom to form a compact structure that interacts with the major groove of DNA in a sequence-specific manner. Generally, in animals, multiple zinc-finger motifs are present as tandem arrays that are separated by a conserved short sequence known as an HC-link (2). These cluster-type zinc-finger proteins interact with contiguous sets of triplet sequences, with each zinc finger making contact with a triplet. In Sp1, Krox20, Zif268, and GAGA, specific amino-acid residues in the $\alpha$-helical region of the DNA-binding surface have been shown to interact with specific nucleotides in target sequences (3–6).

The EPF family is a subfamily of TFIIIA-type zinc-finger proteins of plants. Members of the EPF family have been implicated in floral-organ-specific (7–9) and stress-responsive (10) transcriptional regulation and other regulatory processes (11). The proteins of this family are characterized by the long (19-65 amino acids) linkers of various lengths that separate the zinc fingers. Moreover, the zinc-finger motif itself contains a highly conserved sequence, QALGGH (12, 13), which is located within the region that corresponds to the DNA-contacting surface of TFIIIA-type zinc-finger proteins of animals (3–6).

Our previous DNA-binding studies showed that ZPT2-2, with a linker of 44 amino acids between the two fingers, interacted with two tandemly repeated AGT core sites, which were separated by 10 bp, at a dissociation constant ($K_d$) of 120 nM (14). The binding affinity was sensitive to the spacing between the two core sites. Another protein, ZPT2-1, with a
of these observations and considering the diverse lengths of the linker region among the members of this protein family, we proposed that EPF proteins recognize their cognate DNAs not only by the sequences of the core sites but also by the spacing between the core sites (13, 14).

The binding affinity of ZPT2-2 to the tandemly repeated two AGT sequences is, however, rather weak (Kd=120 nM) compared with those reported for other transcription factors and their cognate binding sequences (Kd = 0.1 to 1.0 nM). This weak affinity could be due to the suboptimal affinities of the respective fingers to the AGT core sequences. These speculations led us to perform SAAB imprinting assays to identify optimal binding sequences of ZPT2-2 from random sequences. We then quantitatively characterized selected sequences by using surface plasmon resonance (SPR) analysis. These experiments have revealed optimal binding sequences for the respective zinc fingers of ZPT2-2 and corroborated our previous conclusion regarding the specificity for the spacing between the two core sites. Moreover, our current results revealed some new features in the protein–DNA interaction of ZPT2-2. These observations allowed us to propose a revised model for the interaction between DNA and ZPT2-2, which is unique among eukaryotic DNA-binding proteins.

Experimental Procedures

**Construction of plasmids.** To construct the plasmids for the production of recombinant proteins in *Escherichia coli*, we used the polymerase chain reaction (PCR) to amplify DNA sequences encoding the N-terminal zinc finger of ZPT2-2 (F1), the C-terminal one (F2), or both by using an upstream primer containing an *Xba*I site at its 5’ end and a downstream primer with a *Pst*I site at its 5’ end. Reaction products were digested with *Xba*I and *Pst*I and were inserted into the pBluescript KS+ (Stratagene, La Jolla, CA, USA) vector to yield pBS-
Unique DNA-interaction of a plant zinc-finger protein

ZPT2-2-F1, -F2, and -F12. To insert a TEV protease site upstream of the coding sequences, synthetic dsDNA for a TEV site with an EcoRV site at its 5' terminus was inserted between the SacI and XbaI sites of pBS-ZPT2-2-F1, -F2, and -F12. These plasmids were digested with EcoRV and PstI, and the resulting fragments were inserted between the XmnI and PstI sites of the expression vector pMAL-c2 (New England Biolabs, Beverly, MA, USA) in-frame with the coding sequence of the maltose binding protein (MBP) to yield pMAL-(TEV)-ZPT2-2-F1, -F2, and -F12.

Expression and purification of truncated forms of ZPT2-2. E. coli BL21 cells were transfected with pMAL-(TEV)-ZPT2-2-F1, -F2 and -F12, and the transformants were grown at 37 °C in LB medium with ampicillin (0.1 mg/mL). After bacterial growth reached an OD$_{600}$ of 0.3 to 0.6, the culture was supplied with isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and allowed to grow for additional 3 h. The cells were harvested, resuspended in buffer A (20 mM Tris-HCl [pH 8.0], 0.01 mM ZnSO$_4$, 2 mM dithiothreitol [DTT], and 0.02% Tween 20), and sonicated with a Sonifier 250 (Branson, Danbury, CT). The resulting cell lysate was centrifuged at 5000 rpm for 15 min, and the supernatant was loaded onto an amylose resin affinity column (New England Biolabs). After the column was washed with buffer A, bound proteins were eluted by using buffer A containing 10 mM maltose. The eluates then were subjected to Hitrap Q (Amersham Pharmacia, Piscataway, NJ, USA) column chromatography with a linear gradient of 0 to 1.0 M KCl in buffer B (20 mM Tris-HCl [pH 8.0] and 0.01 mM ZnSO$_4$). Pooled fractions of fusion proteins were dialyzed against buffer A and subjected to proteolytic digestion with TEV protease (Gibco BRL, Rockville, MD, USA) at 16 °C for 10 h. The reaction mixture was loaded onto an SP column (Waters, Milford, MA, USA) and eluted with a linear gradient of 0 to 1.0 M KCl in buffer B. The fractions of zinc finger proteins were further purified on a C-8 reverse-phase column (Waters) by using a 0% to 100% gradient of acetonitrile in 0.1% (v/v)
trifluoroacetic acid. The purified proteins then were lyophilized, dissolved in 20 mM DTT (pH 10), and dialyzed against a buffer containing 25 mM HEPES-KOH (pH 7.6) and 0.01 mM ZnCl₂ to reconstitute the zinc atoms in the proteins.

**DNA-binding analyses.** Gel-shift assays were performed as follows. Purified proteins were incubated with 5000 cpm of probes (HR5 and HR6), which were labeled with P³² using T4 polynucleotide kinase, at room temperature for 30 min in a binding buffer containing 25 mM HEPES-KOH (pH 7.6), 40 mM KCl, 0.1% Nonidet P-40, 0.01 mM ZnCl₂, and 1 mM DTT. After the reaction, the mixtures were run on an 8% polyacrylamide gel in 0.5 × TBE buffer to separate DNA–protein complexes from excess probe. Hydroxyl radical footprinting was performed as described elsewhere (14) by using the [³²P]end-labeled double-stranded DNA probes shown:

HR5: 5'-CTAGACCGGTGTGGGCCCACACTGGGTGCCCCAACATGGAACGCTAAACCCGGG-TTGGGC-3'

HR6: 5'-GATCCCGCTTGGGCCCACACTGGGTGCCCCAACATGGAACGCTAAACCCGGG-CCG-3'

**Identification of binding sequence.** An optimal binding sequence for ZPT2-2 was identified by using the SAAB imprinting assay (15, 16). Purified P₅₁₂ proteins were incubated at room temperature for 30 min with a mixture of [³²P]end-labeled, double-stranded, 55-bp oligonucleotides containing a central stretch of random sequence (25 bp) and unique sequences of 15 bp at each end for primer-binding sites. The mixtures were subjected to gel shift assays as described above, and DNA bands were made visible by using autoradiography. DNA–protein complexes were eluted from the gel by using 0.5 M ammonium acetate, and the DNA was amplified by 30 cycles of PCR in the presence of [³²P]labeled nucleotides. These procedures of gel shift assays and PCR were repeated for several rounds.
ZPT2-2 and DNA was performed by using the BIAcore biosensor (BIAcore AB, Uppsala, Sweden) based on SPRs that are related to the number of molecules bound to the DNA on the sensor surface. Double-stranded oligonucleotides biotinylated at their 5’ ends were immobilized on a streptavidin-preimmobilized sensor chip (SA; BIACore AB). The sequences of the oligonucleotides were as follows (core sites are underlined):

**AGC:** 5’-CGCCAGAAGCGTCATTT-3’

**CAGT:** 5’-ATTTGACAGTACACGCC-3’

**CAGT-2-AGC:** 5’-GGCCGTCACTTTGACAGTAAAGCGTCATTTTGAGGCCG-3’

**CAGT-6-AGC:** 5’-GGCCGTCACTTTGACAGTACAAAGCGTCATTTTGAGGCCG-3’

**CAGT-8-AGC:** 5’-GGCCGTCACTTTGACAGTACACGAGCGTCATTTTGAGGCCG-3’

**CAGT-10-AGC(G):** 5’-GGCCGTCACTTTGACAGTACACGCGGCAGGGGTACATTTTGAGGCCG-3’

**CAGT-12-AGC:** 5’-GGCCGTCACTTTGACAGTACACGCGCGGCAGGGGTACATTTTGAGGCCG-3’

**CAGT-14-AGC:** 5’-GGCCGTCACTTTGACAGTACACGCGCGCGGCAGGGGTACATTTTGAGGCCG-3’

**CAGT-18-AGC:** 5’-GGCCGTCACTTTGACAGTACACGCGCGCGCGCGAGGGGTACATTTTGAGGCCG-3’

**AGC-3-CAGT:** 5’-GGCCGTCACTTTGAAAGCGATCAGTGTCATTTTGAGGCCG-3’

**AGC-5-CAGT:** 5’-GGCCGTCACTTTGAAAGCGACATCAGTGTCATTTTGAGGCCG-3’

**AGC-7-CAGT:** 5’-GGCCGTCACTTTGAAAGCGACCCATCAGTGTCATTTTGAGGCCG-3’

**AGC-9-CAGT:** 5’-GGCCGTCACTTTGAAAGCGACCCACATCAGTGTCATTTTGAGGCCG-3’

**CAGT-10-TGCT:** 5’-GGCCGTCACTTTGAAAGCGACCCACATCAGTGTCATTTTGAGGCCG-3’

**GCT-7-CAGT:** 5’-GGCCGTCACTTTGGGCTGACCCATCAGTGTCATTTTGAGGCCG-3’

**AGC-7-CAGT-10-AGC:** 5’-GGCCGTCACTTTGAAAGCGACCCACATCAGTGTCATTTTGAGGCCG-3’

**CAGT-10-AGCT:** 5’-GGCCGTCACTTTGAAAGCGACCCACATCAGTGTCATTTTGAGGCCG-3’

**AGCT-7-CAGT:** 5’-GGCCGTCACTTTGGGCTGACCCATCAGTGTCATTTTGAGGCCG-3’
Unique DNA-interaction of a plant zinc-finger protein

PROM: 5'-ATAGATAGAAGCTCAGTCGTAGCCAAAAGCTTTACCATAT-3'

These DNAs were applied to the surface of a sensor chip for a contact time of 5 min, resulting in the capture of 400 to 1000 response units (RU) of the oligonucleotides. Each binding cycle was performed by injecting ZPT2-2 proteins into a constant flow (10 µL/min) of HBPZ (25 mM HEPES-KOH [pH 7.6], 150 mM KCl, 0.01 mM ZnCl₂, and 0.02% Tween 20) buffer across the surface via a sample loop. The sensor chip then was washed with HBPZ buffer for an additional 500 to 1000 s to measure dissociation rates. All reactions were carried out at 25 °C. Data were collected at 2 Hz and analyzed by using the BIAEvaluation program 2.1 (BIAcore AB). We assumed that the reaction between immobilized DNA and ZPT2-2 follows the first-order kinetics model in the following equation:

\[ k_{on} \]

\[ ZPT2-2 + DNA \leftrightarrow ZPT2-2-DNA, \]

\[ k_{off} \]

where \( k_{on} \) is the apparent association rate constant, and \( k_{off} \) is the dissociation rate constant.
Results

Identification of optimal ZPT2-2 binding sequences. To reveal the optimal binding sequences for ZPT2-2, a truncated form of ZPT2-2 (P\textsubscript{F12}) that included both the N-terminal (F1) and C-terminal (F2) zinc fingers was expressed in *E. coli* (Figure 1A), purified to homogeneity (Figure 1B), and subjected to SAAB imprinting assays. The P\textsubscript{F12} protein was incubated with various end-labeled 55-bp oligonucleotides, each of which included a 25-bp central stretch of randomized sequence and binding sites for PCR primers at both ends. The P\textsubscript{F12}-bound oligonucleotides were separated from unbound oligonucleotides by using gel electrophoresis (Figure 2A), eluted from the gel, amplified by PCR, and subjected to further rounds of selection by SAAB imprinting procedure.

Enrichment of ZPT2-2 binding sequences was tested by determining a dissociation constant ($K\text{d}$) for a pool of selected oligonucleotides. The $K\text{d}$ of the interaction between P\textsubscript{F12} and the oligonucleotides after 5 rounds of selection was similar to that after 4 rounds ($K\text{d} = 2.9$ nM; Figures 2B and 2C), which indicated saturation of the selection. Hence, we amplified, cloned, and sequenced the oligonucleotides from the fourth round of selection. A sequence of CAGT was found in all clones (Figure 2D), which suggests that this sequence is the binding site for one of the two zinc fingers. This result is consistent with our previous findings demonstrating that the core binding sites of ZPT2-2 are AGT (14). Further, this current result indicates that a cytosine just upstream of the AGT also is recognized during binding.

Our previous results led us to expect another AGT approximately 10 bp upstream or downstream of the CAGT sequence (14). Unexpectedly, however, such a sequence was not found in any of the clones; instead, an AGC or AGG was found within the randomized region or primer-binding sequences at various distances (either upstream, downstream, or both) from the CAGT sequences. The sequences could be classified into three categories: type I sequences, AGC(T) present at 10 to 12 bp downstream of the CAGT site [CAGT-10~12-
Fig. 1
**type I**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Kd (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ggatccCAACCTCAGCTACGTACACGCCAagctttctgga</code></td>
<td>E1 19</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E9 4.5</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E10 6.2</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E11 9.3</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E13 5.4</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E18 7.5</td>
</tr>
<tr>
<td><code>ggatccCAACCTCAGCTACGTACACGCCAagctttctgga</code></td>
<td>E1 19</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E9 4.5</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E10 6.2</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E11 9.3</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E13 5.4</td>
</tr>
<tr>
<td><code>ggatccCAGCTTAGAAGCTCAGTAACCTACACCCCagctttctgga</code></td>
<td>E18 7.5</td>
</tr>
</tbody>
</table>

**type II**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Kd (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ccTACACACACACACACACACGTCC</code></td>
<td>E3 110</td>
</tr>
<tr>
<td><code>ccTACACACACACACACACACGTCC</code></td>
<td>E4 87</td>
</tr>
<tr>
<td><code>ccTACACACACACACACACACGTCC</code></td>
<td>E8 84</td>
</tr>
<tr>
<td><code>ccTACACACACACACACACACGTCC</code></td>
<td>E15 53</td>
</tr>
<tr>
<td><code>ccTACACACACACACACACACGTCC</code></td>
<td>E12 34</td>
</tr>
</tbody>
</table>

**type III**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Kd (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>gatccACCAATTGAGCGACCCTATCAGTCCCCaagctttctgagggc</code></td>
<td>E6 42</td>
</tr>
<tr>
<td><code>gatccACCAATTGAGCGACCCTATCAGTCCCCaagctttctgagggc</code></td>
<td>E7 31</td>
</tr>
<tr>
<td><code>gatccACCAATTGAGCGACCCTATCAGTCCCCaagctttctgagggc</code></td>
<td>E14 36</td>
</tr>
<tr>
<td><code>gatccACCAATTGAGCGACCCTATCAGTCCCCaagctttctgagggc</code></td>
<td>E17 150</td>
</tr>
</tbody>
</table>

Fig. 2
Unique DNA-interaction of a plant zinc-finger protein

AGC(T)]; type II, AGG present at 10 to 12 bp downstream of the CAGT site (CAGT-10~12-AGG); and type III, AGC present at both 4 to 6 bp downstream and 7 or 8 bp upstream of the CAGT site. Separated by 7 or 8 bp from the CAGT site, an upstream AGC(T) also was present in most of the type I sequences. These results suggest that one of the two zinc fingers of ZPT2-2 recognizes the CAGT, and the other finger has affinity for AGC(T) or AGG.

Binding affinities to the selected sequences as determined by gel shift assays are shown in Figure 2D. The $K_d$ values for type I sequences (0.45 to 1.9 nM) were much lower than those for type II sequences (3.4 to 11 nM). This suggests that ZPT2-2 prefers AGC to AGG for its binding site.

Interaction of the ZPT2-2 zinc fingers with DNA. To characterize the roles of the respective zinc fingers of ZPT2-2 in its interaction with DNA, recombinant proteins containing only a single zinc finger were prepared ($P_{F1}$ and $P_{F2}$, Figure 1) and examined for DNA-binding activity. Because the gel shift assay was not sensitive enough to detect the weak binding activity of the single zinc finger, we instead used an SPR technique. $P_{F1}$ showed weak binding to the AGC site ($K_d = 150 \pm 60 \mu M$; Figure 3A and Table I), and its binding to the CAGT site was undetectable. In comparison, $P_{F2}$ bound to both the CAGT site ($K_d = 4.7 \pm 0.1 \mu M$) and the AGC site ($K_d = 78 \pm 20 \mu M$; Figure 3B and Table I). Therefore, F1 can bind only to the AGC site, whereas F2 can bind to both sites, and its affinity for the CAGT site is much higher than that for AGC. When a probe containing both the core sites (CAGT-10-AGC) was tested, the $K_d$ values were 120 ± 40 µM for $P_{F1}$, 3.9 ± 0.4 µM for $P_{F2}$, and 14 ± 1 nM for $P_{F12}$ (Figure 3D and Table I). Therefore, the binding affinity of the two zinc fingers together is two to three orders of magnitude higher than that of either single finger. This indicates that the two sites in $P_{F12}$ act cooperatively upon binding to DNA. Hydroxyl radical footprinting experiments using $P_{F12}$ revealed footprints at the CATG and the AGCT sites in a
Fig. 3
probe containing CAGT-10-AGCT sequence (Figure 4), indicating that ZPT2-2 indeed interacted with both binding sites.

In light of our current understanding, we propose here a model for the process by which ZPT2-2 binds to its target DNA sequence:

\[
\begin{align*}
K_d &= K_{1} \cdot K_{2} = K_{1}' \cdot K_{2}' \\
1/K_d &= K_1 \cdot K_2 = K_1' \cdot K_2' 
\end{align*}
\]

In eq. 1, \( P_{F12}(F1)\text{--DNA}^+ \) represents an intermediate complex formed by the interaction of the first zinc finger (F1) with the AGC site only; \( P_{F12}(F2)\text{--DNA}^+ \) is an intermediate complex formed by the interaction of the second zinc finger (F2) with the CAGT site only; \( P_{F12}\text{--DNA} \) is a stable complex in which both zinc fingers are bound. \( K_1 \) is an affinity constant for the first interaction of F1 to the AGC site; this constant can be approximated by the \( K_d \) for the binding of \( P_{F1}\text{--AGC} \) (Table I). \( K_2 \) is an equilibrium constant for the second interaction of F2 to the CAGT site, and this interaction is dependent on the prior binding of F1 to the AGC site. \( K_1' \) is an affinity constant for the first interaction between F2 and the CATG site; this constant can be approximated by the \( K_d \) of \( P_{F2}\text{--CATG} \). \( K_2' \) is an equilibrium constant for the second interaction by F1 to the AGC site, and this interaction is dependent on the prior binding of F2 to the CAGT site. As described previously, the affinity of F2 for CAGT is more than 30-fold higher than that of F1 for AGC; this difference is due to both an increased association rate \( (k_{on}) \) and a decreased dissociation rate \( (k_{off}) \) (Table I). Therefore, \( P_{F12}\text{--DNA}_{F2}^+ \) represents a major intermediate, and the binding reaction can be approximated by the two equilibrium constants \( K_1' \) and \( K_2' \). In light of these considerations, we conclude that, upon the
preferentially binding to the CAGT site; this action is followed by the binding of F1 to the AGC site. Although F1’s interaction with the AGC site is relatively weak, it helps to stabilize the ZPT2-2-DNA complex when the F1–AGC interaction occurs in conjunction with the binding of F2 to the CAGT site.

**ZPT2-2 can recognize various arrangements of the two core sites.** In some sequences that were selected by the SAAB imprinting procedure, an AGC(T) sequence is present upstream of the CAGT site in addition to the downstream AGC(T) (Figure 2D). The sequence AGC(T) can be regarded as two overlapping AGCs on two different strands. To examine whether these upstream AGC(T) sequences serve as binding sites, a model probe that contains an AGC sequence 7 bp upstream of the CAGT site (AGC-7-CAGT) was tested for its binding to PF12. The AGC-7-CAGT probe bound to PF12 at a $K_d$ of 44 ± 3 nM (Table I), which indicates that the upstream AGC(T) sequences serve as binding sites for ZPT2-2. Hydroxyl radical footprinting experiments revealed a footprint at this upstream AGCT along with that of the CATG site (Figure 4). This indicates that ZPT2-2 indeed interacted with both binding sites in this configuration, and F1 interacted with the AGC(T) site.

In the sequences selected by the SAAB imprinting procedure, the nucleotide T frequently followed the AGC. We used model probes to examine the molecular basis underlying this sequence preference. PF12 showed stronger binding to CAGT-10-AGCT than to CAGT-10-AGCG (Figure 5). The same preference (AGCT-7-CAGT > AGC-7-CAGT, Figure 5) was observed when the F1-binding site was located upstream of the F2-binding site. Therefore, PF12 prefers T at the position following AGC, regardless of which side of the F2-binding site (CATG) this F1-binding site may be located on.

When F1-binding sites (AGC) were placed on both sides of the F2-binding site (i.e.,
AGC-7-CAGT but lower than that for CAGT-10-AGCG. These results indicate that the upstream AGC acted in an inhibitory manner against the binding of F1 to the downstream AGC (Figure 5).

The differences in the affinities of P_{F12} toward these probes presenting various arrangements of the F1- and F2-binding sites were due primarily to differences in $k_{off}$ values, not to changes in $k_{on}$ values (Figure 5). When P_{F2} was tested for its binding to these probes, only very small differences were observed between all probes. These results indicate that $K_2'$ but not $K_1'$ (eq. 1) was influenced by differences in the arrangement and number of the binding sites.

**Effects of spacing between the core sites.** The P_{F12}-binding sequences selected by the SAAB imprinting procedure (Figure 2D) suggested that ZPT2-2 has a preference regarding the spacing between the F1- and F2-binding sites, in agreement with our previous observations (8, 14). To estimate quantitatively the effects of the spacing, SPR assays were completed by using probes with various spacings between the two binding sites. When the F1-binding site was located downstream of the F2-binding site on the same strand, the highest affinity to P_{F12} was observed for CAGT-10-AGC ($K_d=14 \pm 1 \text{ nM}$; Figure 6A). Increasing the spacing by 2 bp resulted in a 3.5-fold reduction in the binding affinity, and increasing it by 4 bp led to a 5.4-fold reduction. In comparison, decreasing the spacing by 2 bp reduced the binding affinity by 1.8-fold, and decreasing it by 4 bp reduced the binding affinity by 2.4-fold. The reduction in the affinity was due primarily to the increase in $k_{off}$ rather than to the decrease in $k_{on}$ (Figure 6A).

When we tested P_{F2} with the same set of probes, differences in the spacing barely affected the affinity (Figure 6A), which indicates that the spacing does not influence $K_1'$ (eq. 1) during the interaction between ZPT2-2 and DNA. Therefore, the effect of spacing on the
overall binding affinity of ZPT2-2 ($K_{\text{net}}$) is due to the differences in the $K_2$s. These results indicate that the effects of spacing become manifest only upon the second interaction, that is, upon the binding of F1 to the AGC site after the initial binding of F2 to the CAGT site.

We then assessed the effects of spacing for other arrangements of the core sites. In the AGC-n-CAGT arrangement, the binding affinity was again sensitive to the spacing, with the highest affinity being for AGC-5-CAGT ($K_d = 25 \pm 3$ nM, Figure 6B). The effects of spacing on the binding affinity were again due primarily to effects on $k_{\text{off}}$ rather than on $k_{\text{on}}$, and the binding of F2 was not affected by the spacing ($K_d = 20$ to 40 µM, Figures 6A and 6B). Therefore, these results indicate that spacing affects $K_2$ but not $K_1$.

A high-affinity ZPT2-2-binding site is present in the promoter of the ZPT2-2 gene. In light of the DNA-binding specificity of ZPT2-2, we searched for ZPT2-2-binding sites in the upstream region of the ZPT2-2 gene. This action revealed a sequence, CAGTCGTAGCCAAAAGCT (PROM; core sites underlined), 320 bp upstream from a transcription initiation site of ZPT2-2. The PROM sequence contains optimal binding sites for the two zinc fingers with optimal spacing. An SPR assay revealed that $P_{F12}$ bound to PROM with high affinity ($K_d = 9.5 \pm 1.2$ nM, Figure 5). This result strongly suggests that this sequence is a cognate binding site for ZPT2-2 and presumably serves as a cis-element for the autoregulation of the ZPT2-2 gene itself.

**Discussion**

**Proposed model for DNA interaction by ZPT2-2.** In light of our present and previous characterization of the DNA interaction, we here propose a model for the DNA binding of ZPT2-2. ZPT2-2 can interact with two types of DNA sequences, which have different arrangements of the two core sites. The interaction of the F2 zinc finger with the CAGT site
primarily initiates the binding. Then the F1 region interacts with an AGC(T) site that is located several nucleotides either upstream or downstream from the CAGT site, thereby stabilizing the complex. The equilibrium between the intermediate and final complexes greatly depends on the spacing between the CAGT and AGC(T) sites; this spacing determines the overall binding affinity. The unique manner in which ZPT2-2 interacts with DNA is due largely to the flexibility of the long linker region between the two zinc fingers.

Most transcription factors, including cluster-type zinc-finger proteins, interact with continuous DNA sequences. However, some DNA-binding proteins, such as those belonging to the Oct-1 (17, 18) and Zn$_2$Cys$_6$ (19–21) families, are known to recognize separate DNA-sequence motifs, with each of two DNA-binding domains making contact with a different motif. These DNA binding proteins have very strict specificities for the spacing between core sequences, to the degree that even a 1-bp deviation in spacing drastically reduces the DNA binding affinity of Oct-1 (18) and PUT3 (a Zn$_2$Cys$_6$ protein; 21). Compared with those of Oct-1 and PUT3, the specificity of ZPT2-2 for spacing is moderate. In addition, ZPT2-2 will bind to two completely different arrangements of the two DNA-binding sites. Thus, the mechanism by which ZPT2-2 recognizes these noncontiguous binding sites differs from that of Oct-1 and PUT3.

**Biological significance of ZPT2-2’s unique mode of DNA interaction.** Is there any biological advantage to interacting with target DNA sequences in two different arrangements? At present, we have no answer to this question. However, it would be interesting if, depending on various conditions, intact ZPT2-2 could change its conformation between the conformations corresponding to the two arrangements of the two core sites. Because of steric hindrance, the intact ZPT2-2 may not be able to change its conformation as freely as does its truncated form. Even so, the conformation of intact ZPT2-2 may change depending on its
state of protein modification or interaction with other regulatory proteins. Thus, ZPT2-2 could switch from one set of target genes to another in response to environmental or intracellular stimuli.

In the promoter region of the ZPT2-2 gene, we found a putative ZPT2-2-interacting site that is perfectly in accord with the results of our current in vitro study. Many transcription factors are known to regulate the transcription of their own genes through interacting with their promoter DNA sequences (autoregulation). Considering the ubiquity of autoregulation of transcription factors together with the high-affinity interaction of ZPT2-2 with DNA, the sequence we identified most likely serves as a bona fide cis-element for ZPT2-2 in vivo.

Therefore, knowledge of ZPT2-2’s DNA-binding specificity in vitro led us to identifying an in vivo ZPT2-2 interaction site. We expect that this information will facilitate our in vivo functional characterization of ZPT2-2 in the stress response in plants.

Acknowledgements

We thank Dr. Sanjay Kapoor for critically reading this manuscript. This work was supported in part by the COE Promotion Fund from the Science and Technology Agency of Japan and by a PROBRAIN grant from the Bio-Oriented Technology Research Advancement Institution (BRAIN) of Japan.
References


Table I. Kinetic constants for the binding of truncated ZPT2-2 proteins as determined by using the SPR technique. Hyphens represent “data not determined”.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Core sequence</th>
<th>$k_{on}$ (M$^{-1}$s$^{-1}$)</th>
<th>$k_{off}$ (s$^{-1}$)</th>
<th>$K_d$ (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>CAGT</td>
<td>–</td>
<td>–</td>
<td>$\geq 4.0 \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>AGC</td>
<td>$(2.2 \pm 0.5) \times 10^3$</td>
<td>$(3.3 \pm 2.0) \times 10^{-1}$</td>
<td>$(1.5 \pm 0.6) \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>CAGT-10-AGC</td>
<td>$(2.9 \pm 0.6) \times 10^3$</td>
<td>$(3.5 \pm 1.8) \times 10^{-1}$</td>
<td>$(1.2 \pm 0.4) \times 10^{-4}$</td>
</tr>
<tr>
<td>F2</td>
<td>CAGT</td>
<td>$(1.4 \pm 0.1) \times 10^4$</td>
<td>$(6.6 \pm 0.1) \times 10^{-2}$</td>
<td>$(4.7 \pm 0.1) \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>AGC</td>
<td>$(3.0 \pm 0.3) \times 10^3$</td>
<td>$(2.0 \pm 0.3) \times 10^{-1}$</td>
<td>$(7.8 \pm 2.0) \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td>CAGT-10-AGC</td>
<td>$(2.4 \pm 0.3) \times 10^4$</td>
<td>$(9.4 \pm 0.8) \times 10^{-2}$</td>
<td>$(3.9 \pm 0.4) \times 10^{-6}$</td>
</tr>
<tr>
<td>F12</td>
<td>CAGT</td>
<td>$(2.6 \pm 0.4) \times 10^4$</td>
<td>$(5.8 \pm 0.3) \times 10^{-2}$</td>
<td>$(2.3 \pm 0.3) \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>AGC</td>
<td>$(2.2 \pm 0.7) \times 10^3$</td>
<td>$(1.6 \pm 0.1) \times 10^{-1}$</td>
<td>$(7.9 \pm 1.8) \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td>CAGT-10-AGC</td>
<td>$(1.1 \pm 0.2) \times 10^6$</td>
<td>$(1.5 \pm 0.1) \times 10^{-2}$</td>
<td>$(1.4 \pm 0.1) \times 10^{-8}$</td>
</tr>
<tr>
<td></td>
<td>CAGT-11-GCT</td>
<td>$(1.1 \pm 0.2) \times 10^6$</td>
<td>$(5.9 \pm 0.1) \times 10^{-2}$</td>
<td>$(5.3 \pm 0.1) \times 10^{-8}$</td>
</tr>
<tr>
<td></td>
<td>AGC-7-CAGT</td>
<td>$(7.0 \pm 0.5) \times 10^5$</td>
<td>$(3.1 \pm 0.1) \times 10^{-2}$</td>
<td>$(4.4 \pm 0.3) \times 10^{-8}$</td>
</tr>
<tr>
<td></td>
<td>GCT-7-CAGT</td>
<td>$(6.5 \pm 0.1) \times 10^5$</td>
<td>$(2.5 \pm 0.1) \times 10^{-2}$</td>
<td>$(4.0 \pm 0.9) \times 10^{-8}$</td>
</tr>
</tbody>
</table>

Values shown represent the mean ± standard deviation of three experiments.
**Figure Legends**

Figure 1. Production of truncated forms of ZPT2-2 proteins.

A, Schematic representation of full-length ZPT2-2 and its truncated forms. Closed box, basic amino acid cluster (B-box); shaded box, Ser–Thr-rich region; hatched box, zinc-finger motif. Amino acids are numbered in regard to those of the full-length protein. B, SDS-polyacrylamide electrophoresis of purified ZPT2-2 proteins. Lane 1, molecular-weight markers; lane 2, P<sub>F1</sub>; lane 3, P<sub>F2</sub>; lane 4, P<sub>F12</sub>.

Figure 2. SAAB imprint assay.

Gel-shift assays from the (A) first and (B) fourth rounds of selection for optimal ZPT2-2 binding sequences. Concentrations of P<sub>F12</sub> are shown at the top of lanes. The bands indicated by arrowheads were recovered for further processes of SAAB imprinting experiments as shown in “Experimental Procedures”. Bands indicated by open circles are due presumably to 1:2 DNA:protein complexes. C, Binding profile of P<sub>F12</sub> with a mixture of oligonucleotides after fourth-round selection, from which the apparent K<sub>d</sub> value was calculated to be 2.9 nM. D, Alignment of selected ZPT2-2 binding sequences. Bold, binding site (CAGT) for F2; underlining, binding site (AGC and AGG) for F1; double underlining, F1 binding site (AGCT); lower-case letters, primer binding site. The K<sub>d</sub> values determined by the gel shift assays are shown at the right of the figure.

Figure 3. DNA binding specificities of the two zinc fingers. Shown are the sensorgrams of the SPR measurements representing the binding profiles of (A) P<sub>F1</sub>, (B) P<sub>F2</sub>, and (C and D) P<sub>F12</sub> to the CAGT and AGC probes containing a single core site (A, B, and C) or to the CAGT-10-AGC probe containing the two core sites (D). Dosages of proteins are (A) 40 µM P<sub>F1</sub>, (B) 20
μM \( P_{F2} \), (C) 10 μM \( P_{F12} \), and (D) 2.5, 5.0, 10, 20, and 40 nM \( P_{F12} \).

Figure 4. Hydroxyl radical footprinting of ZPT2-2 binding sites.

A DNA fragment containing ZPT2-2 binding sites was labeled with \([\alpha - ^{32}P]dCTP\) at either terminus (\( BamH I \) or \( XbaI \) site) by using the Klenow fragment of DNA polymerase. The probes for top and bottom strands were incubated with (B) or without (F) truncated ZPT2-2 proteins (14), subjected to a hydroxyl radical reaction, and separated on a 15% polyacrylamide–urea sequencing gel with Maxam–Gilbert G sequencing reactions (G). Shown below are densitograms of the banding patterns.

Figure 5. ZPT2-2 can recognize various arrangements of the two core sites.

Binding affinities were measured by the SPR technique and are represented by \( k_{on} \), \( k_{off} \), and \( \Delta G \) values. Schematic representation of probes is shown at the bottom. Closed bar, \( \Delta G \) value for \( P_{F12} \); hatched bar, \( \Delta G \) value for \( P_{F2} \). Values shown represent the mean ± 1 standard deviation of three experiments. \( F2 \)-binding sites (CAGT) are underlined.

Figure 6. Effect of spacing between the two core sites on the binding of ZPT2-2.

Binding affinities were measured by using the SPR technique and are represented by association rates (\( k_{on} \)), dissociation rates (\( k_{off} \)), and free-energy changes (\( \Delta G \)). The effects of spacing were examined for the (A) CAGT-n-AGC and (B) AGC-n-CAGT arrangements. Closed bar, \( \Delta G \) [\( \equiv RT \ln(Kd) \)] value for \( P_{F12} \); hatched bar, \( \Delta G \) [\( \equiv RT \ln(Kd) \)] value for \( P_{F2} \). Values shown represent the mean ± 1 standard deviation of three experiments.
The plant zinc-finger protein ZPT2-2 has a unique mode of DNA interaction
Ken-ichi Yoshioka, Setsuko Fukushima, Toshimasa Yamazaki, Michiteru Yoshida and Hiroshi Takatsuji

J. Biol. Chem. published online July 13, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M104079200

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts