Target-Sequence Recognition by Separate-type Cys$_2$/His$_2$ Zinc Finger Proteins in Plants*

(Hiroshi Takatsui†‡ and Takashi Matsumoto¶)

From the Laboratories of †Developmental Biology and ¶Protein Engineering, National Institute of Agrobiological Resources, 2-1-2 Kannondai, Tsukuba, Ibaraki 305, Japan

The EPF family is a group of transcription factors containing canonical Cys$_2$/His$_2$-type zinc finger motifs that were first discovered in plants. These zinc finger proteins are characterized by two zinc fingers that are separated by spacers of various lengths, which are much longer than typical spacers (HC-link) in cluster-type zinc finger proteins. We describe here direct evidence that the two zinc fingers make contact with two tandemly repeated ACT core sequences that are separated by about 13 base pairs, by contrast to the cluster-type zinc finger proteins that bind to contiguous triplet sequences. DNA binding affinities were sensitive to the spaces between the core sequences, and the sensitivity to the spacing was greatly affected by the DNA sequence between the core sequences, with GC-rich sequences endowing much higher specificity than AT-rich sequences. Among the members of the EPF family, EPF1 was less sensitive to the spacing than EPF2–5. These results suggest that EPFs recognize their cognate target DNAs not only by the sequence of the core sites but also by the spacing between the core sites and, moreover, that different members in the EPF family distinguish their specific target genes by reference to these two parameters. This represents a unique type of target-sequence recognition among Cys$_2$/His$_2$-type zinc finger transcription factors. In addition, site-directed mutagenesis studies demonstrated that the two zinc fingers contribute synergistically to the binding to DNA, indicating that both fingers are necessary for the high affinity DNA binding.

The Cys$_2$/His$_2$ zinc finger proteins, first discovered in the transcription factor IIIA of *Xenopus* (1), represent an important class of eucaryotic regulatory proteins. To date, more than 200 different cDNAs have been found to encode the classical zinc finger motif, and many of their products have been shown to play central roles in development and in general transcription factors. Our analyses clearly demonstrated that the EPF zinc finger proteins make contact with two tandem core sequences that are spatially separated. The DNA binding affinity was found to be sensitive to the spacing between the core sequences, and the sensitivity was dependent on the GC content in the spacer region. Our data also suggest that each member of the EPF family has a specific preference for a particular spacing. The recognition of spacings might be a mechanism that augments target-sequence specificity by compensating for the limited number of nucleotides that can be recognized by only two zinc fingers. Site-directed mutagenesis of the EPF protein revealed that the two zinc fingers act synergistically in the binding to DNA in spite of their separated configuration in the protein structure. Our observations suggest a unique mechanism for target-sequence recognition by the EPF family among Cys$_2$/His$_2$-type zinc finger transcription factors.

EXPERIMENTAL PROCEDURES

Construction of Plasmids—A truncated form of the gene for EPF2–5 (residues 66–210) was created for expression in a rabbit reticulocyte

* This work was supported, in part, by the Enhancement of Center of Excellence, Special Coordination Funds for Promoting Science and Technology, Science and Technology Agency, Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed. Tel.: 81-298-38-8383; Fax: 81-298-38-7417; E-mail: takatsui@ss.abr.affrc.go.jp.
pyranosidewas added to 0.3 mM, followed by incubation for 2 more h to threitol, 10
mM [35S]methionine (1000 Ci/mmol; Amersham Corp.). The sizes and pair(s).
mutant derivatives were digested by pBC SK
were transcribed with T3 RNA polymerase from the T3 promoter in the
20, 10 mM 25 mM HEPES-KOH, pH 7.6, 40 mM KCl, 0.1% Nonidet P-40, 0.01 mM
the method of Bradford (14).
were eluted with a linear gradient of NaCl from 50 mM to 1M in the
buffer. The fractions containing EPF2–5ZF-E (0.8–1.0M NaCl) were
loaded onto an amylose-resin affinity column. The column was washed
digestion mixture was loaded onto a Cibacron Blue 3GA column (Sigma)
amino-terminal MBP domain from the fusion protein to leave a trun-
(1%, w/w; New England Biolabs) at 30°C for 16 h, which removed the
lysate system as follows. The DNA fragment was amplified using an upstream primer (GGGCACTAGGCAATGGGAAACTACCCCGGCTC-
AACTGATACAATCT) that contained a restriction site for XhoI and an
initiation codon and a downstream primer (GGGCCCTCGAGTTCACTT-
CCACTATGACCGCCGCTCAGGTT) that contained a restriction site for XhoI and a stop codon. The amplified DNA fragment was di-
gested with XhoI and XhoI, then inserted into the pBC SK– (Stratagene, La Jolla, CA) vector to yield pBC EPF2–5-ZF. Full-length versions of EPF2–5 and EPF1 were also created by essentially the same procedure.
For production of the truncated form of EPF2–5 in Esherichia coli, the same region of the EPF2–5 gene as described above was amplified by polymerase chain reactions to create additional restriction sites, using an upstream primer that contained a Smal site and a down-
stream primer that contained an XhoI site. The products of polymerase chain reaction were digested with Smal and XhoI and then inserted between the XmnI and SauI sites in an expression vector pMAL-c2 (New England Biolabs, Beverly, MA), in-frame with the maltose-binding protein (MBP) gene upstream of the XmnI site to yield pMAL-EPF2–5.
Site-directed Mutagenesis—Mutations were introduced into the cod-
ing region for the EPF2–5-ZF by the recombinant polymerase chain reaction method (13). Amplification by polymerase chain reaction was carried out with plaque-forming unit DNA polymerase (Stratagene) using mutant forward primers and an M13-21 reverse primer for the carboxyl-
terminal portion of the protein, mutant reverse primers complimentary to the mutant forward primer, and an M13 reverse primer for the amino-terminal portion, respectively, with pBC EPF2–5-ZF as the tem-
plate. One-tenth each of the products from the two reactions was com-
bined and the second polymerase chain reaction was carried out with these templates, which were complimentary to each other at one end, using the M13 reverse primer and the M13-21 primer. Reaction prod-
ucts were digested with XhoI and XhoI and then inserted into the pBC SK— vector. All constructs were confirmed by DNA sequencing.
Transcription and Translation In Vitro—pBC EPF2–5-ZF and its mutant derivatives were digested by XhoI, treated with proteinase K, extracted with phenol:chloroform (1:1, v/v), and then purified with a Wizard DNA cleanup kit (Promega Corp., Madison, WI). The templates were transcribed with T7 RNA polymerase from the T7 promoter in the pBC SK– vector, and the transcripts were purified by extraction with phenol:chloroform (1:1, v/v) and ethanol precipitation. One µg of tran-
scripts was translated in a rabbit reticulocyte lysate in the presence of [35S]methionine (1000 Ci/mmol; Amersham Corp.). The sizes and amounts of the products were checked by electrophoresis in a 12.5% SDS-polyacrylamide gel and subsequent autoradiography. Then, equiva-
 lent amounts of the products were used for subsequent gel-shift ex-
periments. The wild-type EPF2–5-ZF protein produced in this system is referred to hereinafter as EPF2–5ZF-R.
Full-length EPF2–5 and EPF1 were also expressed in the rabbit
reticulocyte lysate in a similar manner. The proteins are referred to as
EPF2–5-R and EPF1-R, respectively.
Expression of a Truncated Form of EPF2–5 in E. coli and Its Purification—pMAL-EPF2–5ZF was introduced into E. coli strain TB1. The transformants were grown at 37°C in LB medium with ampicillin (0.1 mg/ml) until A600 reached 0.5; and then isopropyl-β-D-thiogalacto-
pyranoside was added to 0.3 mM, followed by incubation for 2 more h to induce the expression of the MBP-EPF2–5ZF fusion cell.
Products were harvested and lysed by sonication, and the resulting cell lysate was loaded onto an amyllose-resin affinity column. The column was washed with buffer (10 mM sodium phosphate, pH 7.2, 1 mM NaCl, 0.25% Tween 20, 10 mM β-mercaptoethanol, 0.5 mM phenylmethylsulfonyl fluoride, and 10 µM zinc acetate), and the proteins that bound to the column were eluted with the same buffer plus 10 mM maltose. The eluted proteins were concentrated and subjected to proteolytic digestion with factor Xa (1%, w/w; New England Biolabs) at 30°C for 16 h, which removed the amino-terminal MBP domain from the fusion protein to leave a truncated form of EPF2–5 with no junction sequence (EPF2–5ZF-E). The digestion mixture was loaded onto a Cibacron Blue 3GA column (Sigma) that had been equilibrated with 20 mM Tris-HCl, pH 7.5, 1 mM dithio-
reitol, 10 mM zinc acetate, and 50 mM NaCl, and the bound proteins were eluted with a linear gradient of NaCl from 50 mM to 1 M in the buffer. The fractions containing EPF2–5ZF-E (0.8–1.0 M NaCl) were
pooled and concentrated. Protein concentrations were determined by the method of Bradford (14).
DNA Binding Assays—All DNA binding reactions were carried out in 25 mM HEPES-KOH, pH 7.6, 40 mM KCl, 0.1% Nonidet P-40, 0.01 mM

1 The abbreviations used are: MBP, maltose-binding protein; bp, base pairs.

Target-Sequence Recognition by Separate-type ZF Proteins

ZnCl2, 10 µg/ml poly(dI-dC), and 0.1 mM dithiothreitol. Gel-shift assays were performed with 10,000 cpm of end-labeled probe and about 1 µl of the product of in vitro translation. After incubation for 20 min at room temperature, the mixtures were subjected to electrophoresis in a 0.7% agarose/3% polyacrylamide gel, as described previously (11). To determine the dissociation constant for the binding of EPF2–5ZF to its target DNA, we incubated 0.1 µg of unlabeled probe DNA with EPF2–5ZF-E. After electrophoresis in a 10% polyacrylamide gel in 0.5 x TBE buffer (0.09 Tris borate, pH 8.0, 0.002 M EDTA), the gel was stained with ethidium bromide, and then free and bound DNA were quantified by scanning the gel with a CCD camera-based densitometer (ATTO, Tokyo, Japan) under UV light (254 nm).

Hydroxyl radical footprinting was performed as described by Tullius and Dombroski (15) with modifications. A 32P-end-labeled probe (200,000 cpm) was incubated with 5 ng of EPF2–5ZF-E in 30 µl of binding buffer for 20 min. Hydroxyl radical reactions were started by the addition of 2 µl each of 0.45% H2O2, 0.15 mM [(NH4)2Fe(SO4)2]-EDTA, and 50 mM sodium ascorbate, and the reactions were quenched after 2 min by the addition of 5 µl of 0.1 M thiourea. The reaction products were subjected to electrophoresis in a 0.7% agarose/3% poly-
acrylamide gel, and DNA-protein complexes (1:1) were eluted from the gel with 0.5 mM ammonium acetate after autoradiography. The eluate was extracted twice with phenol and once with ether, and DNA was precipitated twice with ethanol. The purified digestion products were separated on a 12% polyacrylamide-urea sequencing gel. Autoradiography was carried out with an image scanner (BAS 2000, Fuji Photo Film, Kanagawa).

RESULTS

EPF2–5 Makes Contact with Two Separate AGT Motifs—The amino acid sequences of the two zinc finger regions in EPF2–5 and EPF1 are shown in Fig. 1. There is substantial sequence similarity between the two zinc fingers of each protein and also between those of the two proteins, in particular around the strongly conserved QALGGH motifs. One of the features of the structures of these proteins is that the two fingers are separated by spacers of different lengths, with 44 and 61 amino acids in EPF2–5 and EPF1, respectively, between the second His in the first finger and the first Cys in the second finger. The products of the genes for both EPF1 and EPF2–5 were origi-
nally found to bind to a tetramer of the EP1 sequence (TGATTTTGACAGTGTCACCTT), the binding sequence of a petal-specific nuclear factor within the promoter region of the EPSPS gene (11, 12). Subsequently, a tetramer of the truncated form of EP1 (EP1S:TTGACAGTGTCAC) was found to have much stronger affinity for these proteins. This observation led us to speculate that the spaces between each unit of the binding sequence in the multimers might be a critical factor in the difference in binding affinities for the two probes and, moreover, that the long spacer region between the two zinc fingers might be responsible for the recognition of the spacing on the DNA.

To characterize the interactions of the EPFs and their target DNAs, we first determined minimal binding sequences by scanning mutation of the target DNA sequences using a truncated form of EPF2–5 that has been expressed in a rabbit reticulocyte in vitro translation system (EPF2–5ZF-E). Assuming from previous observations that each one of the two zinc fingers interacts with one unit of the EP1S, we used two tandemly repeated sequences as probes and competitors. An initial competition experiment with competitors with two-base mutations revealed that mutations in the CAGT sequence severely reduced the competition (M2 and M3), whereas those in the region that flanked the CAGT sequence had little effect (M1, M4, and M5; Fig. 2A). Subsequent experiments with one-base mutations revealed that mutation of any one base in the sequence AGT (M7, M8, and M9) reduced the binding, whereas a mutation in the preceding C (M6) had little effect. Furthermore, substitution of all other bases outside the three bases had little effect on the competition (M18; Fig. 2B). These results indicate that the tandemly repeated AGT is the minimal core sequence for the high affinity binding of EPF2–5.

For further characterization of the protein-DNA complex using larger amounts of the EPF2–5ZF protein, we overexpressed a fusion protein with MBP in E. coli and subsequently released the EPF2–5ZF domain from the MBP domain by proteolytic cleavage. Fig. 3 shows that the protein was purified to near homogeneity. The apparent dissociation constant ($K_d$), determined using EPF2–5ZF-E and a dimer of the EP1S sequence as a probe, was 120 nM (Fig. 4). Although this value indicates that the affinity of interaction is relatively low as compared to other DNA-binding proteins, it is similar to the binding of Tramtrack protein from Drosophila, which contains two clustered zinc fingers, to its native binding site (∼400 nm) (16).

Hydroxyl radical footprinting provides high-resolution details of a protein-DNA complex because of the small size of the DNA-cleavage reagent. We used this technique to obtain additional information about the EPF-DNA complex. A mixture of EPF2–5ZF-E and a probe DNA containing two tandemly repeated EP1S sequences was subjected to the hydroxyl radical in vitro translation system (EPF2–5ZF-E). Assuming from previous observations that each one of the two zinc fingers interacts with one unit of the EP1S, we used two tandemly repeated sequences as probes and competitors. An initial competition experiment with competitors with two-base mutations revealed that mutations in the CAGT sequence severely reduced the competition (M2 and M3), whereas those in the region that flanked the CAGT sequence had little effect (M1, M4, and M5; Fig. 2A). Subsequent experiments with one-base mutations revealed that mutation of any one base in the sequence AGT (M7, M8, and M9) reduced the binding, whereas a mutation in the preceding C (M6) had little effect. Furthermore, substitution of all other bases outside the three bases had little effect on the competition (M18; Fig. 2B). These results indicate that the tandemly repeated AGT is the minimal core sequence for the high affinity binding of EPF2–5.

For further characterization of the protein-DNA complex using larger amounts of the EPF2–5ZF protein, we overexpressed a fusion protein with MBP in E. coli and subsequently released the EPF2–5ZF domain from the MBP domain by proteolytic cleavage. Fig. 3 shows that the protein was purified to near homogeneity. The apparent dissociation constant ($K_d$), determined using EPF2–5ZF-E and a dimer of the EP1S sequence as a probe, was 120 nM (Fig. 4). Although this value indicates that the affinity of interaction is relatively low as compared to other DNA-binding proteins, it is similar to the binding of Tramtrack protein from Drosophila, which contains two clustered zinc fingers, to its native binding site (∼400 nm) (16).

Hydroxyl radical footprinting provides high-resolution details of a protein-DNA complex because of the small size of the DNA-cleavage reagent. We used this technique to obtain additional information about the EPF-DNA complex. A mixture of EPF2–5ZF-E and a probe DNA containing two tandemly repeated EP1S sequences was subjected to the hydroxyl radical reaction, and EPF2–5ZF-DNA complexes (1:1) were purified on and eluted from a gel. A sequencing gel revealed clear footprints at two tandem AGT sequences that were separated by 13 bp in the top strand and complementary sequences (ACT) in the opposite strand (Fig. 5). We had previously excluded the possibility that EPF binds to DNA as a multimer (12), and therefore, our results indicated that one molecule of EPF2–5ZF binds to two separate AGT core sequences, with each independent zinc finger making contact with one AGT sequence. The bands of DNA between the core sequences in the footprint, in particular those adjacent to the AGT core sequences, were somewhat weaker in intensity than those outside the core sequences, suggesting some interaction of the DNA with the spacer region between the two zinc fingers. We have not yet determined the core binding sequences of the other members of the EPF family. However, at least EPF1 and EPF2–4 are very likely to bind to the same AGT core sequences because these two proteins bind to the EP1S tetramer with similar affinity to that of EPF2–5, and the amino acid sequences in the zinc finger motifs of the respective proteins are very similar (12).

**EPF Recognizes the Spacing between Two Core Sequences—**

Next, we examined whether the spacings between the core sequences in the DNA affects the binding affinity. We used probes with different spacings, in which AT or GC repeats were inserted between the EP1S units. The results showed that EPF2–5 was moderately sensitive to the spacing when the spacer region was AT-rich (Fig. 6a). The affinity was highest with a spacing of 13 and 15 bp. However, suboptimal spacings were well tolerated. By contrast, when the spacer sequences were GC-rich, the recognition of the spacing was much stricter than with the AT-rich sequences (Fig. 6b), and the binding affinity showed a sharp peak at a spacing of 13 bp (Fig. 6c). These results indicate that EPF2–5 recognizes the spacing between the two core sequences in the target DNA and that the
sequence between the cores, presumably its AT or GC content, greatly affects the sensitivity of the protein to the spacing. The strict sensitivity to the spacing, in particular with the GC-rich spacers, suggests that the protein structure in the spacer region of the EPF2–5 protein is relatively rigid and that the spatial configuration of the two fingers is relatively stable. The difference in the sensitivity to AT- and GC-rich spacers was presumably due to differences in the flexibility of the DNA in the spacer region. Another possibility is that the direction of DNA bending upon the binding of the EPF protein might be responsible for the difference. In other words, the direction of the DNA bending with AT-rich spacers was more favorable for binding with suboptimal spacing. However, we did not detect any DNA bending with spacers of either base composition (data not shown). Therefore, it appears that it is simply the difference in structural flexibility that is likely to be the cause of the differences in sensitivity.

In the DNA binding experiments in Fig. 6, we used full-length EPF2–5 proteins instead of the truncated proteins that were used in the other experiments. The gel-shift assays indicated that the binding affinities of both forms were similar, which suggest that there is no additional determinant of DNA sequence specificity other than the zinc finger region, at least with the probes used in this work.

Each EPF Protein Has a Different Preference for a Particular Spacing—To compare the recognition of the spacing by different members of the EPF family, the same series of the “spacing-mutant” probes was examined for binding to EPF1. As shown in Fig. 6d, the binding affinity was similarly high with probes from 12 to 16 bp, even with GC-rich spacers. This recognition of the spacing was very relaxed as compared to that by EPF2–5, which showed stronger specificity for a spacing of 13 bp. This difference could be attributable to differences in the number and the sequence of the amino acids in the spacer regions between the two fingers in the two proteins.

Although the difference in the recognition of spacing between EPF1 and EPF2–5 was rather minor, it is quite possible that each member of the EPF family distinguishes its target genes from differences in spacing, as well as from the primary sequences of the three nucleotides recognized by each finger.
compared to that of cluster-type zinc finger proteins. Apparently more independent nature of each zinc finger as suggested by our X-ray studies, abolished the binding almost completely. These results indicate that two fingers are necessary for high affinity DNA binding to EPF2–5–ZF-R. b, effects of disruption of either one of two zinc fingers on DNA binding. The first His residue in either of the two zinc fingers of EPF2–5–ZF-R was replaced by Asp by site-directed mutagenesis, as described under "Experimental Procedures." The mutated proteins carrying a His-to-Asp mutation in the first finger (lane 1) and in the second finger (lane 2), respectively, and wild-type protein (lane 3) were expressed in a rabbit reticulocyte lysate system and subjected to gel-shift assays with a tetramer of the EP1S as a probe. Positions of shifted bands are shown by arrows. The faint bands with lower mobility in some lanes are due to the binding of two protein molecules to the probe.

**Fig. 6.** Recognition of the spacing between the two core sequences in the target DNA. a, probes containing four tandem repeats of the EP1S sequences with various spacings between each unit were tested for binding to EPF2–5–R. The probes contained TA × 2n (n = 1 – 4) sequences between each unit of the EP1S or a two-base deletion from both ends of the EP1S. The spacing between the centers of the two AGT sequences is shown at the top of each lane. The probes were end-labeled and subjected to gel-shift assays with EPF2–5–R. b, probes containing GC × 4n (n = 1–2) sequences instead of AT × 2n in a were tested for binding to EPF2–5–R. c, probes containing spacers of GC repeats differing sequentially by one nucleotide between the two EP1S sequences in a dimer were tested for the binding to EPF2–5–R. d, the same set of the probes as in c was tested for the binding to EPF1. Positions of shifted bands are indicated by arrows. The faint bands with lower mobility in some lanes are due to the binding of two protein molecules to the probe.

**Fig. 7.** Two fingers are necessary for high affinity DNA binding. a, effect of a mutation in one half-site of the tandem repeats in the target DNA. A wild-type probe (lane 1) and a mutant probe with an AGT-to-CGT mutation in the second core sequence in the dimer of the tandem repeats (lane 2) were tested for binding to EPF2–5–ZF-R. b, effects of disruption of either one of two zinc fingers on DNA binding. The first His residue in either of the two zinc fingers of EPF2–5–ZF-R was replaced by Asp by site-directed mutagenesis, as described under "Experimental Procedures." The mutated proteins carrying a His-to-Asp mutation in the first finger (lane 1) and in the second finger (lane 2), respectively, and wild-type protein (lane 3) were expressed in a rabbit reticulocyte lysate system and subjected to gel-shift assays with a tetramer of the EP1S as a probe. Positions of shifted bands are shown by arrows. The faint bands with lower mobility in some lanes are due to the binding of two protein molecules to the probe.

**DISCUSSION**

The structure of EPF zinc finger proteins is unique in that each molecule has two zinc finger motifs that are separated by very different lengths between the two fingers. Two Zinc Fingers Contribute Synergistically to the Binding to DNA—A common perception is that a protein must have a minimum of two fingers for high affinity DNA binding by Cys2/His2-type zinc finger proteins. However, in the EPF family, two zinc fingers appear to make contact with core DNA sequences even more independently than they do in proteins with clustered zinc fingers. To address this issue, we first tested a probe in which one of the two core sequences was tested for the binding to EPF1. Positions of shifted bands are indicated by arrows. The faint bands with lower mobility in some lanes are due to the binding of two protein molecules to the probe.

The structure of EPF zinc finger proteins is unique in that each molecule has two zinc finger motifs that are separated by

---

2 K. Kubo, A. Sakamoto, A. Kobayashi, Z. Rybka, and H. Takatsuji, manuscript in preparation.
type transcription factors that contain Cys_{2}/Cys_{2} zinc finger motifs. The nuclear receptors are known to recognize the spacing in the target sequence, thereby distinguishing cognate promoter sequences from those of other members of the protein family (17, 18). However, these factors bind to DNA as dimers, with each subunit binding to a half-site in the repeats, and this binding is clearly different from that by the EPF family, whose members bind to tandemly repeated sequences as a monomer.

There is another characteristic structural feature of the zinc finger region of the EPF family: strongly conserved sequence (QALGGH motif) within the region that corresponds to the α-helical region in several zinc finger proteins that have previously been characterized. This region, which is strongly conserved in the classical zinc finger motifs of many transcription factors, has been shown to be a DNA recognition surface that faces the target DNA. In a subclass of zinc finger proteins, such as Sp1, Krox 20, and Zif268 (6–10), three positions in the α-helical region have been discovered to participate in major interactions with nucleotides (CXXCXXXFYXXLXYX...‘HXXXXH). These amino acids have been shown to recognize one nucleotide in the triplet of the target sequence by mutational analysis, as a general rule. In the EPF family, the strong conservation of the QALGGH motif in the corresponding region suggests that this region serves as a DNA-recognition surface. Phage display selection from a library of zinc fingers showed that Ala at position b might participate in base recognition (20); however, the second Gly at position c is unlikely to make contact with bases, at least via its side chain, because this amino acid has no side chain. If one or two amino acids in this region participate in the recognition of nucleotides, the fixed sequence of the QALGGH motif in more than 30 proteins limits still further the diversity of the target sequences that can be recognized by the members of the EPF family. Taking into account these factors, we cannot neglect the importance of the length of the spacings.

We are in the process of analyzing the three-dimensional structure of a complex between EPF2–5 and its target DNA by x-ray crystallography and two-dimensional nuclear magnetic resonance. The information from these analyses should reveal the amino acids involved in the recognition of the DNA. We also expect our physicochemical analyses to reveal the role of the spacer region between the two fingers in the contacts with DNA. The optimal spacing of 13 bp between the ATG core sequences for the binding of EPF2–5 is longer than one helical turn (10.5 bp). It will be of interest to determine whether the spacer region lies along the major groove of DNA, as has been proved for cluster-type zinc finger proteins, such as GLI (21), or whether it makes contact with the core sequences across the major groove.

The SUPERMAN gene is an important regulator of development in Arabidopsis thaliana. Recessive mutations of this gene cause extra stamens to form interior to the normal third whorl stamens (22). This gene was recently cloned and found to encode a protein that contained only one classical zinc finger with the QALGGH motif (23). Several other genes (At ZFP family) in Arabidopsis that encode proteins with one zinc finger with the QALGGH motif have also been reported (24). So far, all the Cys_{2}/His_{2}-type zinc finger proteins that have been found in plants have the QALGGH motifs in spite of the different strategies used for cloning their genes, suggesting that single and double zinc finger proteins that contain the QALGGH motif form a major class of transcription factors in plants. Target sequences for these single-finger proteins have not yet been identified, although the AP3 and PI genes are potential candidates for the target genes of the SUPERMAN protein. With respect to DNA binding, our results indicate that one finger recognizes only three nucleotides; therefore, one finger is obviously insufficient for target-sequence specificity. Additionally, we have demonstrated that more than one finger is necessary for high affinity DNA binding. Two possibilities can account for the recognition of target sequence by the single-finger proteins: 1) some other region of the protein is involved in the recognition of DNA, probably in combination with the zinc finger; and 2) the protein binds to DNA as a dimer or in association with other sequence-specific DNA-binding proteins. These possibilities should be taken into consideration when the target sequences are explored. The presence of a basic leucine-zipper motif in these proteins, which might be a domain involved in protein-protein interactions, supports the second possibility.

Acknowledgments—We are grateful to Shin-ichi Omura, Yoshiko Ouchi, Tomiko Yasuhara, and Mayumi Ishikawa for technical assistance.

REFERENCES
1. Miller, J., McLachlan, A. D., and Klug, A. (1985) EMBO J. 4, 1609–1614
2. Rosenberg, U. B., Schroder, C., Preiss, A., Kienlin, A., Cote, S., Riede, I., and Jackle, H. (1986) Nature 319, 336–339
3. Tautz, D., Lehmann, R., Schurch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K., and Jackle, H. (1987) Nature 327, 383–389
4. Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schroder, C., Kemler, R., and Jackle, H. (1986) Cell 47, 1025–1032
5. Nakaseko, Y., Neuhaus, D., Klug, A., and Rhodes, D. (1992) J. Mol. Biol. 228, 619–636
6. Nardelli, J., Gibson, T. J., Vaseque, C., and Charnay, P. (1991) Nature 349, 175–178
7. Nardelli, J., Gibson, T., and Charnay, P. (1992) Nucleic Acids Res. 20, 4137–4144
8. Dejarlais, J. R., and Berg, J. M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7345–7349
9. Dejarlais, J. R., and Berg, J. M. (1992) Proteins 12, 101–104
10. Pavletich, N., and Pabo, C. (1991) Science 252, 809–817
11. Takatsui, H., Mori, M., Benfey, P. N., Ren, L., and Chua, N.-H. (1992) EMBO J. 11, 241–249
12. Takatsui, H., Nakamura, N., and Katsumoto, Y. (1994) Plant Cell 6, 947–958
13. Higuchi, R. (1990) in PCR Protocols: A Guide to Methods and Applications (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eda), pp. 177–183, Academic Press, Inc., San Diego
14. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
15. Tullius, T. D., and Dombroski, B. A. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 5469–5473
16. Fairall, L., Harrison, S. D., Travers, A. A., and Rhodes, D. (1992) J. Mol. Biol. 226, 349–366
17. Naar, A. M., Boutin, J.-M., Lipkin, S. M., Yu, V. C., Holloway, J. M., Glass, C. K., and Rosenfeld, M. G. (1992) EMBO J. 11, 4507–4517
18. Choo, Y., and Klug, A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 11163–11167
19. Pavletich, N. P., and Pabo, C. O. (1993) Science 261, 1701–1707
20. Bowman, J. L., Sakai, H., Jack, T., Weigel, D., Mayer, U., and Meyerowitz, E. M. (1992) Development (Cam.) 114, 599–615
21. Sakai, H., Medrano, L. J., and Meyerowitz, E. M. (1995) Nature 378, 199–203
22. Tague, B. W., and Goodman, H. M. (1995) Plant Mol. Biol. 28, 267–279
Target-Sequence Recognition by Separate-type Cys$_2$/His$_2$ Zinc Finger Proteins in Plants
Hiroshi Takatsuji and Takashi Matsumoto

J. Biol. Chem. 1996, 271:23368-23373.
doi: 10.1074/jbc.271.38.23368

Access the most updated version of this article at http://www.jbc.org/content/271/38/23368

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

This article cites 23 references, 7 of which can be accessed free at http://www.jbc.org/content/271/38/23368.full.html#ref-list-1