*Journal of Plant Biology*, December 1999, 42(4) : 302-309

# Potential Target Gene and Protein Interaction of *Pimpinella brachycarpa* HD-Zip Protein, Phz4 by Yeast Two-Hybrid System

**Sanghyun Lim[1], Yong-Hwan Moon[2], Jung-Il Cho[1], Jong Chan Woo[1], and Kwang-Woong Lee[1]***

[1]Department of Biology, Seoul National University, Seoul 151-742, Korea
[2]Department of Plant and Microbial Biology, University of California, Berkeley, California 94720-3012, USA

**The yeast two-hybrid system was used to investigate dimerization between proteins of *Phz2* and *Phz4* clones of the homeodomain-leucine zipper family which were obtained by screening a *Pimpinella brachycarpa* shoot-tip cDNA library. Assays showed that Phz4 formed a homo rather than a heterodimer with Phz2. In addition, we isolated cDNA clones, *Phyb1*, *Phyb2*, and *Phyb3*, that encode proteins interacting with Phz4. Although Phyb1 is not a HD-Zip protein, the activity of interaction between Phyb1 and Phz4 was, surprisingly, stronger than that of the homodimerization of Phz4. The analysis of interacting parts indicated that from 1 bp to 466 bp of Phyb1, there was no interaction with Phz4, but from 467 bp to 593 bp, interactions were found with the N-terminal and C-terminal regions, except for HD-Zip of Phz4. This region of *Phyb1* contained a nuclear localization signal. DNA-binding analysis showed that the Phz4 HD-Zip domain recognized the [T(C/G)ATTG] core sequence and the region containing the [TCATTG] motif, which is, in itself, a promoter in vitro.**

*Keywords*: homeodomain, leucine zipper, protein-protein interaction, target gene

Homeobox genes were discovered in *Drosophila melanogaster* by Gehring (1987). The DNA-binding domain of these genes, termed homeobox, is characterized by 61 conserved amino acids (Hoey and Levine, 1988; Muller et al., 1988). These genes play important roles in the regulation of developmental processes in man and animals, as well as in yeast. The homeodomain-leucine zipper (HD-Zip) proteins differ from other homeodomain (HD) proteins in that they contain a leucine zipper motif which is closely linked to the carboxy terminal end of the HD region.

Although a large number of homeobox genes have been isolated from various animal species, the HD-Zip motif so far has been found only in plant genes. They were first discovered in *Arabidopsis thaliana* (Ruberti et al., 1991; Mattson et al., 1992; Schena and Davis, 1992; Carabelli et al., 1993), and have also been identified in carrot (Kawahara et al., 1995), tomato (Meissner and Theres, 1995), and maize (Ingram et al., 1999). HD-Zip genes help control cell-fate specification. For example, *Glabra 2* from *Arabidopsis* is required for root-hairless cell identity in roots (Masucci et al., 1996), and is also expressed in all cell layers of young, developing leaves (Szymanski et al., 1998). Another HD-Zip gene, *Kn1* from maize, is required for maintaining meristematic cells in an undifferentiated state (Long et al., 1996). In addition, the *Arabidopsis* gene, *AtML1*, is expressed in the embryonic protoderm from the very early stages of development (Lu et al., 1996).

Leucine zippers are sequence elements that are responsible for dimerization in a separate class of transcription factors found in other eucaryotes. The exact spatial register between the homeodomain and the leucine zipper motif in the Athb-1 and Athb-2 HD-Zip (Ruberti et al., 1991) is similar to that observed between the DNA-binding and the dimerization domains in the b-Zip proteins, another class of transcription factors (Vinson et al., 1989). In the former relationship, a 9-bp dyad-symmetric DNA sequence is bound as a homodimer (Sessa et al., 1993). The ability to form homo or heterodimers between members of the HD-Zip family in vivo could provide enormous potential for generating highly specific mechanisms for the regulation of gene expression. The yeast two-hybrid system, which is applicable to any cloned gene, can be used to detect physical interactions between Phz2 and Phz4 HD-Zip proteins isolated from the *Pimpinella brachycarpa* shoot-tip cDNA library (Moon et al., 1996). This sensitive method can also help identify cDNA clones encoding proteins that interact with Phz4 as a way to study protein-protein interactions.

Study concerning in vivo target genes for plant HD-Zip proteins is extremely important. Target DNA sequences are examined for extensive deletion mutational analysis (Ziyu et al., 1999) and random bind-

*Corresponding author; fax +82-2-872-6881
e-mail kwlee@plaza.snu.ac.kr

ing site selection. Core motifs of HD-Zip are distinct from the [TNATTG] sequences by binding site selection (Sessa et al., 1993). Athb-1 HD-Zip recognizes [T(A/T)ATTG] core sequences, while the high-affinity binding sites of Athb-2 HD-Zip are [T(C/G)ATTG] (Aoyama et al., 1995). Elements similar or identical to this motif are present in numerous other plant gene promoters (Plesch et al., 1997). In fact, *Arabidopsis* PRHA was identified by its ability to specifically bind, in vitro, to an 11-bp motif, C**TAATTG**TTTA, within the promoter of a pathogen-inducible parsley gene, *pr2* (Korfhage et al., 1994). In addition, the *even-skipped* (*eve*), a HD protein in *Drosophila*, appears to control morphogenesis by regulating the expression of the segmentation gene *engrailed* (*en*), and by auto-regulating its own expression (Kawahara et al., 1995).

It was previously shown that Phz2 and Phz4 have high homology to Athb-2 (Ruberti et al., 1991) and THOM1 (Meissner and Theres, 1995) in the HD-Zip motif (Moon et al., 1996). In this report, we describe the protein interaction of Phz4 by yeast two-hybrid system, and a potential target gene for this protein.

## MATERIALS AND METHODS

### Isolation and Sequencing of 5′ Portion of Phz4 cDNA

We isolated the 5′ upstream region of Phz4 cDNA from 5 μg of shoot-tip genomic DNA in *P. brachycarpa*. PCR amplification was then performed using the LA PCR in vitro cloning kit (TaKaRa). The PCR was performed in a thermal cycler (Pharmacia) for 33 amplification cycles. Each cycle comprised of 95°C for 1 min, 57°C for 1 min, and 72°C for 2 min, all in the presence of the C1 and GENES1 primers (Table 1). The amplified products were purified and reamplified using C2 and GENS2 primers that contained restriction sites for cloning. The BamHI-XbaI fragment generated by PCR was subcloned into Bluescript SK II, using standard techniques (Sambrook et al., 1989). This fragment was then deleted using a kilo-sequencing deletion kit (TaKaRa). The deleted DNA series were sequenced by the dideoxy chain-termination technique (Sanger et al., 1977) using a sequenase ver. 2.0 (United States Biochemical) or a thermo sequenase cycle sequencing kit (Amersham), according to manufacturers protocols.

### Construction of Fusion Protein

We used PCR to construct plasmids for yeast two-hybrid system that were flanked by EcoRI and SalI sites. The PCR products from the cDNA templates of Phz4 and Phz2 were ligated into the EcoRI and SalI sites of pADGAL4 and pBDGAL4. To ascertain the degrees of interaction between them, four parts of Phz4 and three parts of Phyb1 were prepared as shown in Figure 1. The PCR products cut with the EcoRI and SalI sites were ligated into pADGAL4 and pBDGAL4. For Phz4 HD-Zip expression, the HD-Zip domain of Phz4 was subcloned into pFLAGATS of XhoI and BglII. Recombinant FALG (FLAG:HD-Zip) fusion protein was expressed in *Escherichia coli* and purified by affinity chromatography that contained FLAG resin (Sigma). Protein concentration was determined by using the method of Bradford (1976).

### Yeast Two-Hybrid Assays

Yeast two-hybrid assays were performed with strain

**Table 1.** Cassettes and primers used for PCR amplification of the 5′ upstream region of the *Phz4* genomic DNA.

| | |
|---|---|
| BamHI Cassette | |
| | 5′ HO GTACATATTGTCGTTAGAACGCGTAATACGACTCACTATAGGGA 3′ |
| | 3′ CATGTATAACAGCAATCTTGCGCATTATGCTGAGTGATATCCCTCTAG OH 5′ |
| HindIII Cassette | |
| | 5′ HO GTACATATTGTCGTTAGAACGCGTAATACGACTCACTATAGGGAGA 3′ |
| | 3′ CATGTATAACAGCAATCTTGCGCATTATGCTGAGTGATATCCCTCTTCGA OH 5′ |
| XbaI Cassette | |
| | 5′ HO GTACATATTGTCGTTAGAACGCGTAATACGACTCACTATAGGGAGAT 3′ |
| | 3′ CATGTATAACAGCAATCTTGCGCATTATGCTGAGTGATATCCCTCTAGATC OH 5′ |
| Cassette Primer C1 | 5′ GTACATATTGTCGTTAGAACGCGTAATACGACTCA 3′ |
| Cassette Primer C2 | 5′ CGTTAGAACGCGTAATACGACTCACTATAGGGAGA 3′ |
| GENS1 | 5′ ATTAAAAGCGACGCGGTAGTGACTAGAACGCGA 3′ |
| GENS2 | 5′ GGTTGGCATGGAAGGTGCTAGATTTAGTTGGAG 3′ |

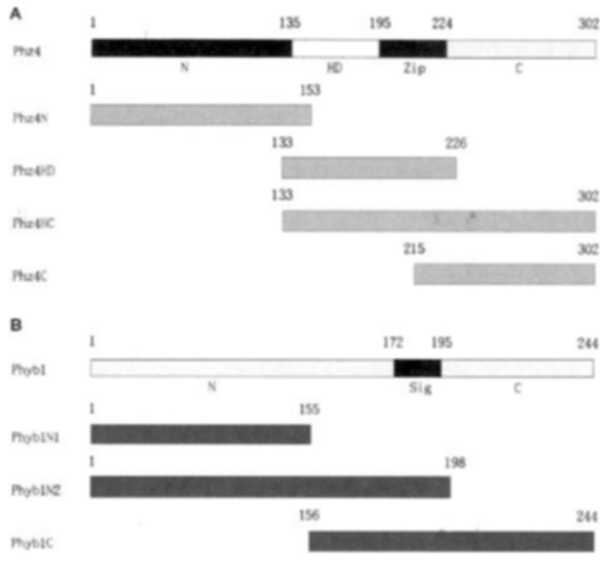



**Figure 1.** The schematic representation of primers used in amplification of *Phz4* cDNA fragments (A) and *Phyb1* cDNA fragments (B). N, N-terminal; C, C-terminal; HD, Homeodomain-leucine; Zip, leucine zipper; Sig, a potential nuclear localization signal.

YRG-2 (Mata, *ura3-52*, *his3-200*, *ade2-101*, *lys2-801*, *trp1-901*, *leu2-3*, *112*, *gal4-542*, *gal80-538*, LYS::$UAS_{GAL1}$-$TATA_{GAL1}$-HIS3, URA3::$USA_{GAL417mer(x3)}$-$TATA_{CYC1}$-*lacZ*). Double transformants were grown at 28°C for two to three days on solid minimal medium that lacked histidine, leucine, and tryptophan. Protein interaction was defined as the expression of a *lacZ* reporter gene in a colony filter assay. Positive colonies that appeared blue were quantified by an *o*-nitrophenyl-β-D-galactoside (ONPG) assay. The activity unit was calculated using the formula: 1000 $OD_{420}/[OD_{600} \times$ reaction time (min) × volume of culture (mL)]. We performed yeast transformations and β-galactosidase assays using X-gal and ONPG as a substrate, according to the Clontech Two-Hybrid System II Instruction Manual.

## Electrophoretic Mobility Shift Assay (EMSA)

The protein extracts were prepared from *E. coli* DH5α as follows: LB medium with 100 μg $mL^{-1}$ ampicillin was inoculated with a 1/100 dilution of an overnight culture, then shaken at 37°C. After an $OD_{600}$ of 0.8 was reached, 0.5 mM IPTG was added, and incubation was continued for 8 h. Afterward, cells were collected by centrifugation and resuspended in 10 mL of a buffer containing 10 mM Tris-HCl (pH 7.5) and 150 mM NaCl. Cells were sonicated at 75% maximal setting for five rounds of 30 s each, with intermediate resting intervals of 30 s. Unbroken cells and cell debris were removed by centrifugation. The purification of Phz4 was performed according to the Kodak manual. EMSA reactions contained 20 ng purified Phz4, 3 ng $^{32}$P-end labeled probe, 150 ng competitor DNA, and 2 μg polyd(I-C). The buffer for these reactions comprised 20 mM Hepes, 50 mM KCl, 2 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, 0.5% NP-40, and 10% glycerol. The reaction mixture was incubated for 20 min at room temperature. It was then loaded onto 5% acrylamide/bisacrylamide (37.5:1) gels in 0.25 × TBE at 20 mA for 2 h. After the gels were dried, they were autoradiographed using Kodak films.

# RESULTS

## Screening cDNA Clones Encoding Proteins Interacting with Phz4

To identify proteins that interact with Phz4, a truncated Phz4 containing the homeodomain-leucine zipper and C-terminal region (HC) was fused to a GAL4 DNA-binding domain. A reconstitution GAL4 activity might be any protein encoded by an activation domain that includes total cDNA and that interacts with Phz4. The colonies that grew on a medium lacking histidine, leucine, and tryptophan were examined for β-galactosidase activity (Table 2). After total DNA was transformed into *E. coli*, we obtained the cDNA clones, *Phyb1*, *Phyb2*, and *Phyb3*, that were ligated in pADGAL4. One clone, *Phyb2*, was a partial *Phz4*. Another *Phyb3*, differed from *Phz4* in the sequence below the leucine zipper (Fig. 2). Interestingly, the other clone, *Phyb1*, was not in the HD-Zip

**Table 2.** Quantitative assay of dimer formation.

| Activation domain plasmid | Binding domain plasmid | β-Galactosidase activity[a] |
|---|---|---|
| Phyb1 | 4HC | 7.46 ± 0.16 |
| Phyb2 | 4HC | 1.52 ± 0.07 |
| Phyb3 | 4HC | 0.33 ± 0.02 |
| 4HD | 4HD | 0.24 ± 0.02 |
| 4HC | 4HD | 0.30 ± 0.02 |
| 4HC | 4HC | 4.37 ± 0.11 |
| 2HC | 2HC | 3.64 ± 0.09 |
| 4HC | 2HC | 0.08 ± 0.01 |

4, Phz4; 2, Phz2; HD, homeodomain-leucine zipper; HC, homeodomain-leucine zipper including C-terminal.
[a]β-Galactosidase activity unit = 1000 × $OD_{420}/[OD_{600}$ reaction time (min) × volume of culture (mL)].

```
Phz4    G P S S T P V E T P R P H H S G S S H H R V A F N P W A I A  290
Phyb3   - I - - - L T - - L - - Q R L K R V V P I R L R T L L T L A  284

Phz4    P A G H R S F D A V P H                                      302
Phyb3   S Q L K I I I W L L G Y                                     296
```

**Figure 2.** The aligned amino acid sequences of the C-terminal region of *Phz4* and *Phyb3*.

```
G TTT CTC CAA ATC AAG ATG GCG TCG TCA TTG GGT TCG CTG AAA TCG GTG ATT TTC GAT   58
                M  A  S  S  L  G  S  L  K  S  V  I  F  D                      14

CGC CAA GCA AGA AAA CAG CAA TAT CAA TCT CAT ATA TTA GGC CTC AAT GCT TAC GAT CGC  118
R  Q  A  R  K  Q  Q  Y  Q  S  H  I  L  G  L  N  A  Y  D  R                    34

CAC AAA AAA TTC ATC AGT GAT TAT GTT GAC TTC TAT GGG AAA GAT GTG TCA GCA GAA GAA  178
H  K  K  F  I  S  D  Y  V  D  F  Y  G  K  D  V  S  A  E  E                    54

AAG CTG CCT GTT AAA ACT GAT AAA GAT ACT CTT AGA GAA GGT TAT CGA TTT ATA AGA TCC  238
K  L  P  V  K  T  D  K  D  T  L  R  E  G  Y  R  F  I  R  S                    74

GAG GAA GAT GAT CTG GAT CCC TCC TGG GAG CAG AGA CTA GTG AAG CGC TAC TAT AAT AAG  298
E  E  D  D  L  D  P  S  W  E  Q  R  L  V  K  R  Y  Y  N  K                    94

CTT TTT AAA GAA TAC TGT ATA GCT GAT ATG TCA CAA TAC AGG ACA GGC AAG ATT GGC CTC  358
L  F  K  E  Y  C  I  A  D  M  S  Q  Y  R  T  G  K  I  G  L                   114

AGA TGG AGA ACA GAA AAA GAA GTA ATA TCT GGA AAA GGG CAG TTT GTA TGC GGT AAT AAA  418
R  W  R  T  E  K  E  V  I  S  G  K  G  Q  F  V  C  G  N  K                   134

CAT TGT GAT GAA AAA GAT GGC CTT GCA AGC TAC GAG GTA AAC TTT TCT TAT TTT GAG GCC  478
H  C  D  E  K  D  G  L  A  S  Y  E  V  N  F  S  Y  F  E  A                   154

GGG GAG AGC AAA CAA GCC CTT GTG AAA TTA GTA ACC TGT GAG AGA TGT GCA AAG AAG CTG  538
G  E  S  K  Q  A  L  V  K  L  V  T  C  E  R  C  A  K  K  L                   174

AAT TAC AAA AAA CAG AAA GAG AAA TTA AAA AGA AAT GAG AAA GAG GTG CTT CGA AGG AAA  598
N  Y  K  K  Q  K  E  K  L  K  R  N  E  K  E  V  L  R  R  K                   194

AGG GAA AGG TCA GAA AGT GAT GAT GAT ATT CCT ACT GAG TAC AAG GGG AAC AAA GAC AGC  658
R  E  R  S  E  S  D  D  D  I  P  T  E  Y  K  G  N  K  D  S                   214
AGG AAA GTC ATG AAG ACC TCC ACT TCC TTG GGA GAC CAG AAG GCT GGT GAT ACT GAA AAC  718
R  K  V  M  K  T  S  T  S  L  G  D  Q  K  A  G  D  T  E  N                   234

TTT GAT GAG TTT CTT GAG GGA ATG TTT CTT TAA GTT TGA GTC AAC GAT GAC AGC TGC AAT  778
F  D  E  F  L  E  G  M  F  L  *                                              244

ATG CTA TCT GCC TTG AGA TGA TCA TTT CCC CCC GGG TAG TGT GAA TTA TAT ATA GTG TAG  838
TTG TAA CTG GAA TAG TCA GAA TCT CCT AAT GCT TAT CTG TGA CAA ATT TCA TTC TGT CGT  898
AAT AAC TAT TGG ACT AGA GTA ATT TTA TAT TTT AAG GTA ATT TAT TTT CCT TGT ATA ATG  958
GCG TTA GAC GAG TTT AAA CTT TGG GTT GTA ATT CTT TTT TCT TTG AGT TTC AGA GTT TTA 1018
AAT GAA AAA AAA AAA AAA AAA AA                                               1041
```

**Figure 3.** The entire nucleotide sequence and the deduced amino acid sequence of the *Phyb1* cDNA. The amino acid sequences underlined indicate a potential nuclear localization signal. Stop codon is marked with an asterisk. The GenBank accession number of this sequence is AF082024.

family. The primary structure of Phyb1 was deduced from the sequence of a 1041-bp region containing an open reading frame of 725 codons (Fig. 3).

## Parts Interacting with Phz4 and Phyb1

The activity of Phyb1 with Phz4 was five times stronger than that of Phyb2, a partial *Phz4* (Table 2). To determine the regions responsible for this strong interaction between Phz4 and Phyb1, we divided the *Phz4* cDNA into four parts (Fig. 1A), and *Phyb1* cDNA was deleted from the 5' or 3' ends to create truncated fusion proteins (Fig. 1B). When 1 and 155

**Table 3.** Degree of dimerization between various regions of Phyb1 and Phz4.

| Activation domain plasmid | Binding domain plasmid | β-Galactosidase activity[a] |
|---|---|---|
| Phyb1N1 | 4N | 0.03 ± 0.006 |
| Phyb1N1 | 4HD | 0.02 ± 0.005 |
| Phyb1N1 | 4HC | 0.03 ± 0.005 |
| Phyb1N1 | 4C | 0.04 ± 0.008 |
| Phyb1N2 | 4N | 3.79 ± 0.101 |
| Phyb1N2 | 4HD | 0.08 ± 0.010 |
| Phyb1N2 | 4HC | 0.10 ± 0.011 |
| Phyb1N2 | 4C | 4.37 ± 0.113 |
| Phyb1C | 4N | 4.79 ± 0.121 |
| Phyb1C | 4HD | 3.26 ± 0.987 |
| Phyb1C | 4HC | 7.35 ± 0.204 |
| Phyb1C | 4C | 9.78 ± 0.342 |

N, N-terminal; C, C-terminal; HD, homeodomain-leucine zipper; HC, homeodomain-leucine zipper including C-terminal. [a]β-Galactosidase activity unit = 1000 $OD_{420}/[OD_{600}$ × reaction time (min) × volume of culture (mL).

amino acids of Phyb1 (Phyb1N1)/pAD fusions formed a complex with the four parts (4N, 4HD, 4HC, 4C) of the Phz4/pBD fusions, β-galactosidase activities were almost zero (Table 3). The combinations of 156 and 244 amino acids of Phyb1 (Phyb1C), however, apparently had activities. The 1 and 198 amino acids of Phyb1 (Phyb1N2) contained a nuclear localization signal, and appeared to have β-galactosidase activity in regions that included the N-terminal (4N) and the C-terminal (4C) of Phz4. This led us to conclude that the N- and C-terminal regions of a Phz4 are indispensable for interaction with a Phyb1, and that the interacting part of Phyb1 includes a nuclear localization signal (Fig. 3).

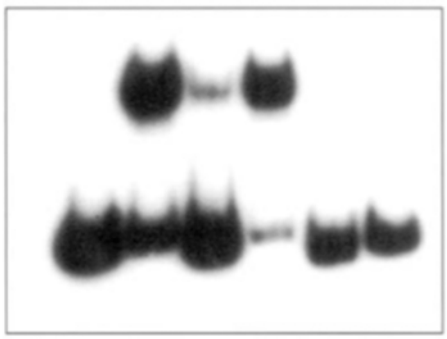
## Dimer Formation of Phz2 and Phz4

We amplified two regions of HD and HC to determine the formation of homo or heterodimers of Phz2 and Phz4, the member of the HD-Zip family isolated from the *P. brachycarpa* shoot-tip cDNA library. Both regions were ligated into pAD/pBDGAL4, and transformed in yeast. The colonies that grew on a medium lacking histidine, leucine, and tryptophan were exam-

ined for β-galactosidase activity (Table 2). These analyses showed that the C-terminal sequence below the leucine zipper greatly affected the degree of dimerization, which suggests that the interaction of Phz2 and Phz4 (including the HC region) preferred homodimerization to heterodimerization. This result was the same as for the degree of interaction about Phyb2, a partial Phz4, and Phyb3, the other HD-Zip cDNA with Phz4. Therefore, we concluded that homodimer formation of Phz4 is favored over heterodimer formation.

A

| Oligo | I | I | I | II | II | V |
|---|---|---|---|---|---|---|
| Competitor | × | × | ○ | × | ○ | × |
| Protein | × | ○ | ○ | ○ | ○ | ○ |



B

| Oligo | III | III | III | IV | IV | V |
|---|---|---|---|---|---|---|
| Competitor | × | × | ○ | × | ○ | × |
| Protein | × | ○ | ○ | ○ | ○ | ○ |

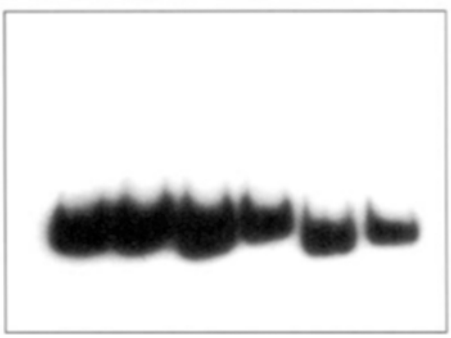


**Figure 4.** Phz4 binds different sequences with different affinities by EMSA. (A) Oligo I, II and V motifs probes are [T**C**ATTG], [T**G**ATTG] and [T**CG**TTG], respectively. (B) Oligo III and IV motifs probes are [T**A**ATTG] and [T**T**ATTG], respectively. Each probe (3 μg) was incubated with a partial Phz4 HD-Zip purification protein (20 μg); the competitive DNA probe used 150 μg. The only vector besides the Phz4 protein was used as a negative control.

## DNA-Binding Properties of Phz4 in Vitro

For studying DNA-binding properties, we used the binding sites of the *Arabidopsis* Athb-1 and Athb-2 HD-Zip proteins, referred to as [T(A/T)ATTG] and [T(C/G)ATTG], respectively. These were determined by selecting high-affinity binding sites from random sequence DNA (Sessa et al., 1993). Using EMSA, we analyzed DNA-binding properties for the [TNATTG] core sequence, known as HD-Zip recognition sites with partial Phz4. To assess the importance of the core sequence, we also exchanged the third base A for G. Phz4 HD-Zip bound the [T(C/G)ATTG] core sites and showed no detectable interaction with the [T(A/T)ATTG] binding sites and with the [TCGTTG] binding sites (Fig. 4, A and B). A 50-fold molar excess of unlabelled [T(C/G)ATTG] fragment competed effectively for probe binding, indicating the specificity of the retarded complex. Therefore, the DNA-binding sites of Phz4 were the [T(C/G)ATTG], as would be found in members of the HD-Zip II family (Moon et al., 1996).

## Potential Target Sites of Phz4

The early embryo *even-skipped* (*eve*) in *Drosophila* appears to control morphogenesis by regulating the expression of the segmentation gene *engrailed* (*en*),

```
GCATAACTCAAAGTTGACTTAGCAGATCTTCGATAATATTTCTTTGTCCTTGTCTTGCCTCTTTCTGCATACTGGACCC  -1078
GTTGTTTCATCAAGAAATCTTTGAATAGCGTCTACAGCTTCATTCTTCTCTGGAGTGCCGCACCCTATACTTCTCAAGTT  -999
ACCTTTCAACCTTGATCAATTCACTTCTTCCATACATTATCAGACTTGCCTTATCCATGTAAAAACTCTTGAATATATCA  -919
TTGCTTATGTACTTATCTATTGGAGGCACTGAGCTAATTTTTGGACTTGCCTTGATCAACGAGGTCTTTTAGGGTTTCAA  -839
TAAAGTAACTGATTGAGCTTAGAGTTCCTTGGAACTTTAGTCATAAAATACCTTTAGCTCTGTGAAAGACCTTTTATAGAT  -759
AGAACTTGGTTGAGATTGTGAACATTCCATGACCTACAAGAGATTTTGAAATCTGCCATGAATTTTCAGGTCTCTGGCCC  -679
AACCTTCAGATCTTCGATAGTATTTCATAGAGAAGTGAAGTATAATTAAATGTTTGTATTTATTACTTTGAGATCATAGG  -599
AGTATCAATTCCACTTCTGAATCATCATTATGTTCTAGTCAAGCCTGATGTCCATTAAACATTTTGATGGCATCCGACTG  -519
TAACTCCCATTCAGATCCTTTCTTTAGTGCAACATCCACTTTATTATTTTCTCTTTCTTCATTATACTTTTTGAATCTAG  -439
CAGATCTTCGATAGCAGATCTTATCCCTAGGTGGTAACTTTAAATCGGCATAGCAGCTTAGATCCTGGAAAAACGTTCCA  -359
TAGCAATACGATTAGGCCATAATAATGATCTAGGATCAATGGCTTTTAAAGCCGAATCGATCGGCAAATATCCGATCCGA  -279
ACCATCAGGACCATTGCTAGCTTACGCCATCTAGATATAGCTACGAGACTGCTAGCTACGGCTTAAAAATCGCTTAGCTA  -199
TACGACTCGGATCGACGGCTAGAATCGGTTTCGAAAAATCGGTTTCGGGGCTATTATCGGGCTAAAATCGCAGCAGCTGA  -119
ATCGATCGGCATATTTCTATCGCCTATCTCGATTTCGATATAATACCTCGATCCTAGCCCTAGGATTCCAAAAGCTTCGA  -39
TGTATAAATGTCCTTTCGCTTCTTAGCTTACTATCTCGGAAGAACATATATATTTTTCAAGAAAGGCTAGGCTAACCACC  42
CCATAGTAACCTAAGCTTGTTCCTTGCATATAAATCAAGAAAAGCTAGTTTATATCAAGAATTTATATATACATATAAGC  122
ATG ATG ATT CAT CAA AGA GAA GAT CAC TTG GGT CTA AGC CTA AGC TTG TCC TCT CCA GCT  182
 M   M   I   H   Q   R   E   D   H   L   G   L   S   L   S   L   S   S   P   A   20
```

**Figure 5.** The 5′ upstream region of the *Phz4* nucleotide. Underlined Italic type indicates the primer sequences used in amplification of DNA fragments for EMSA with Phz4.

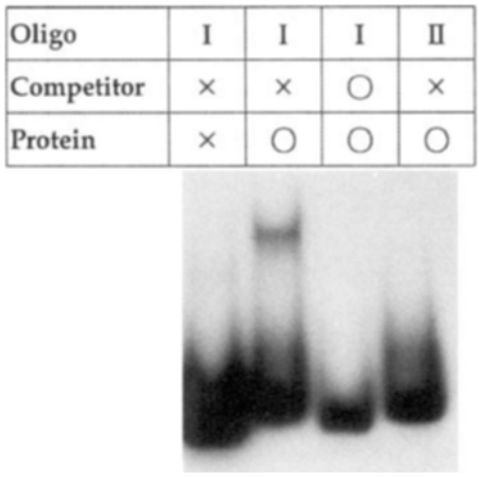



**Figure 6.** Electrophoretic mobility shift assay of Phz4 HD-Zip motif protein binding to the DNA target sites. Oligo I is the TA**TCATTG**CTTA motif; oligo II is the TT**TCATAG**AGAA motif.

and by autoregulating its own expression (Hoey and Levine, 1988). To learn whether Phz4 HD-Zip protein binds to its own promoter, as the HD proteins do, we investigated the 5′ upstream region of Phz4 (Fig. 5). We succeeded in obtaining a PCR product using the BamHI, HindIII, and XbaI cassette primers and GENES1. Using the BamHI cassette primer and GENS2, we obtained a secondary PCR product. In addition, the 5′ upstream sequence analysis of Phz4 showed that the Phz4 promoter, itself, contained the [TCATTG] at positions –921 bp to –916 bp (Fig. 5). It also contained a [TCATAG] sequence similar to the [TCATTG]. We tested whether the *Phz4* promoter region encompassing these motifs could be bound by Phz4. In EMSA experiments, Phz4 was bound to the Phz4 promoter region containing the [TCATTG] motif, but was not bound to the [TCATAG] motif (Fig. 6). Altering the core [TCATTG] sequence to [TCATAG] resulted in a loss of binding ability. When unlabelled homologous DNA was increased by 50-fold molar amounts, the strength of binding was decreased. We concluded that a potential target site of Phz4 HD-Zip protein is, in itself, a promoter in vitro.

## DISCUSSION

### Protein-Protein Interaction of Phz4

We examined the protein interaction of the Phz4 by using yeast two-hybrid system. Because leucine zippers were involved in protein-protein interactions, we anticipated that the Phz4 HD-Zip protein would act as a dimer. Whereas the β-galactosidase activity of Phz4HD and Phz4HD was only 0.24, that of Phz4HC and Phz4HC was 4.37 (Table 2). We presumed that the C-terminal of the Phz4 HD-Zip protein was related to form a dimer. The b-zip, Fos, one of the transcription factors was unable to homodimerize, but readily formed heterodimers with another b-zip protein, Jun (Lewin, 1991). Evidently the formation of homo or heterodimers between members of the HD-Zip families can increase the number of DNA bindings and widen the potential for distinct activities in controlling gene expression.

Study of protein dimerization has established that Athb-1 and Athb-2 can form homodimeric complexes in vitro (Sessa et al., 1993). The high homology of two leucine zippers that were identical in 35 out of 39 positions suggests that they would probably allow for the formation of homodimers although they are included in different groups. In addition, Hahb-1 within the HD-Zip II subfamily and Hahb-10 in the I subfamily have formed homodimers (Gonzalez et al., 1997). In our study, Phz2 and Phz4 in the same subfamily preferentially formed homodimers (Table 2).

Because understanding the interaction among HD-Zip proteins has become more critical, we isolated three cDNA clones. These clones encoded proteins Phyb1, Phyb2, and Phyb3 that interacted with Phz4 from the *P. brachycarpa* shoot-tip pADGAL library. Interaction activity about Phyb2, a partial Phz4 cDNA, with Phz4 was approximately five times stronger than that of the interaction about Phyb3, another HD-Zip protein, with Phz4 (Table 2). The interaction of Phz4 and Phz4 also was stronger than that of Phz4 and Phz2. Based on these two results, we believe it is possible that Phz4 acts as a homodimer preferentially.

Surprisingly, Phyb1, which was isolated by the two-hybrid system using Phz4 as bait, was not a member of HD-Zip family (Fig. 3). Nevertheless, the activity of interaction was stronger than that of the homodimer complex with Phz4. The *Phyb1* cDNA clone included a nuclear localization signal, and the RNase protection assay revealed that *Phz4* and *Phyb1* were expressed at similar rates in all organs, including leaves, petioles, roots, and shoot tips (Moon, 1997). The characterization of Phyb1 will be helpful for studying the function of the Phz4HD-Zip, a transcription factor.

### Target Sites of Phz4

We focused on identifying the natural target genes of Phz4 because it is important to know the action of

HD-Zip proteins in plant development. Phz4 bound to the T(C/G)ATTG motif. This sequence was distinct from the binding sites CAAT(A/T)ATTG and CAAT(C/G)ATTG. These sites were shown to interact with the *Arabidopsis* HD-Zip proteins Athb-1 and Athb-2, respectively, in vitro, through binding site selection (Sessa et al., 1993; Aoyama et al., 1995). In vitro and in vivo DNA-binding studies revealed that Oshox-1 from rice interacted with the pseudopalindromic sequence CAAT(C/G)ATTG (Meijer et al., 1997). Likewise, Gmh1 from soybean was bound to the CAAT(C/G)ATTG (Moon, 1997). In our study, Athb-2, Gmh1, Oshox1, and Phz4, all members of the HD-Zip II family, bound to the CAAT(C/G)ATTG (Fig. 4). These results demonstrate that grouping of the HD-Zip family could be done according to amino acid sequences, as well as by DNA-binding sites.

It is important to seek target genes when identifying binding sites. PRHA from *Arabidopsis* was reported as having the capacity to bind to TAATTG core sequence elements of the parsley pathogenesis-related gene, *pr2*, but it required additionally adjacent bases for high-affinity binding (Korfhage et al., 1994; Plesch et al., 1997). In our study, Phz4 was identified as binding to the TCATTG motif within the promoter itself (Fig. 5). It is likely that Phz4 autoregulated its own expression (Fig. 6). Future research addressing the characterization of *Phyb1* cDNA and discovery of natural target genes will provide insight into the function of the HD-Zip protein for signal transduction in plants.

## ACKNOWLEDGEMENTS

This study was supported by a grant from the Seoul National University Research Fund (97-06-2080) to K.-W. Lee.

Received December 6, 1999; accepted December 10, 1999.

## LITERATURE CITED

Aoyama T, Dong CH, Wu Y, Carabelli M, Sessa G, Ruberti I, Morelli G, Chua NH (1995) Ectopic expression of the Arabidopsis transcriptional activator Athb-1 alters leaf cell fate in tobacco. Plant Cell 7: 1773-1785

Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye binding. Anal Biochem 72: 248-254

Carabelli M, Sessa G, Baima S, Ruberti I (1993) The *Arabidopsis Athb-2* and *-4* genes are strongly regulated by far-red-rich light. Plant J 4: 469-479

Gehring WJ (1987) Homeoboxes in the study of development. Science 236: 1245-1252

Gonzalez, DH, Valle EM, Chan GGRL (1997) Interaction between proteins containing homeodomains associated to leucine zippers from sunflower. Biochim Biophys Acta 1351: 137-149

Hoey T, Levine M (1988) Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. Nature 332: 858-861

Ingram GC, Magnard JL, Vergne P, Dumas C, Rogowsky PM (1999) ZmOCL1, an HDGL2 family homeobox gene, is expressed in the outer cell layer throughout maize development. Plant Mol Biol 40: 343-354

Kawahara R, Komamine A, Fukuda H (1995) Isolation and characterization of homeobox-containing genes of carrot. Plant Mol Biol 27: 155-164

Korfhage U, Trezzni GF, Meier I, Hahlbrock K, Somssich IE (1994) Plant homeodomain protein involved in transcriptional regulation of a pathogen defense-related gene. Plant Cell 6: 695-708

Lewin B (1991) Oncogenic conversion by regulatory changes in transcription factors. Cell 64: 303-312

Long JA, Moan EL, Medford JI, Barton MK (1996) A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. Nature 379: 66-69

Lu P, Porat R, Nadeau JA, ONeill SD (1996) Identification of a meristem L1-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox gene. Plant Cell 8: 2155-2168

Masucci JD, Rerie WG, Foreman DR, Zhang M, Galway ME, Marks MD, Schiefelbein (1996) The homeobox gene *GLABRA2* is required for position-dependent cell differentiation in the root epidermis of *Arabidopsis thaliana*. Development 122: 1253-1260

Mattson J, Soderman E, Svenson M, Borkird C, Engstrom P (1992) A new homeobox-leucine zipper gene from *Arabidopsis thaliana*. Plant Mol Biol 18: 1019-1022

Meijer AH, Scarpella E, Dijk EL, Qin L, Taal AJC, Rueb S, Harrington SE, McCouch SR, Schilperoort RA, Hoge JH (1997) Transcriptional repression by Oshox1, a novel homeodomain leucine zipper protein from rice. Plant J 11: 263-276

Meissner R, Theres K (1995) Isolation and characterization of tomato homeobox gene THOM1. Planta 195: 541-547

Moon YH (1997) Molecular cloning and characterization of homeodomain-leucine zipper genes in *Glycine max* and *Pimpinella brachycarpa*. Ph.D. thesis. Seoul National University, Seoul, pp 83-85

Moon YH, Choi SB, Kim JI, Kim JC, Han TJ, Cho SH, Lee KW (1996) Isolation and characterization of three homeodomain-leucine zipper genes, *Phz1*, *Phz2*, and *Phz4* from Pimpinella brachycarpa shoot tips. Mol Cells 6: 697-703

Muller MM, Ruppert S, Schaffner W, Matthias P (1988) A cloned octamer transcription factor from lymphoid-specific promoters in non-B cells. Nature 336: 544-551

Plesch G, Stormann K, Torres JT, Walden R, Somssich IE (1997) Developmental and auxin-induced expression of the *Arabidopsis prha* homeobox gene. Plant J 12: 635-647

Ruberti I, Sessa G, Lucchetti S, Morelli G (1991) A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. EMBO J 10: 1789-1791

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, 2nd ed, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463-5467

Schena M, Davis RW (1992) HD-Zip proteins: Members of an *Arabidopsis* homeodomain protein superfamily. Proc Natl Acad Sci USA 89: 3894-3898

Sessa G, Morelli G, Ruberti I (1993) The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. EMBO J 12: 3507-3517

Szymanski DB, Jilk RA, Pollock SM, Marks MD (1998) Control of GL2 expression in *Arabidopsis* leaves and trichomes. Development 125: 1161-1171

Vinson CR, Sigler PB, Mcknight SL (1989) Scissor-grip model for DNA recognition by a family of leucine zipper proteins. Science 246: 911-916

Ziyu D, An K, Edward GE, An G (1999) Functional role of CAAT box element of the nopaline synthase (*nos*) promoter. J Plant Biol 42: 181-185