Functional Analysis of Homeodomain-Leucine Zipper Transcription Factors in Arabidopsis thaliana

BY

HENRIK JOHANNESSON
Dissertation for the Degree of Doctor of Philosophy in Physiological Botany presented at Uppsala University in 2000

Abstract


Homeodomain-leucine zipper (HDZip) proteins constitute a large family of transcription factors apparently unique to plants. To elucidate the function of these factors, the biochemical properties in vitro as well as the effects on transgenic plants when expressed at high levels were studied. The conclusion is that HDZip proteins are very similar with respect to DNA-binding specificity in vitro but appear to be active in different aspects of plant development. Thus, functional specificity of HDZip proteins is most likely determined by other aspects of proteins function, e.g. their capacity to interact with other proteins.

High-level expression of the HDZip gene ATHB5 in transgenic plants results in hypersensitivity to the inhibitory effect of the plant hormone abscisic acid (ABA) on seed germination and seedling root growth. Furthermore, the expression of ATHB5 in germinating seedlings is downregulated in the Arabidopsis ABA response mutants abi3 and abi5. Together, these data suggest that ATHB5 acts as a regulator of seed germination and postgerminative growth downstream of ABI3 and ABI5 in an ABA response-signaling pathway.

Enhanced levels of the HDZip gene ATHB13 in transgenic Arabidopsis confer a sugar-dependent reduction of cotyledon width. In addition, a subset of known sugar-dependent genes was hyperinduced by sucrose in Arabidopsis seedlings overexpressing ATHB13. These data suggest that ATHB13 affects both cotyledon morphology and gene regulation as a component of a sucrose-signaling pathway.

Loss-of-function mutations in ATHB5 and ATHB13 did not result in any discernable mutant phenotypes, suggesting that these genes are only required under specific physiological conditions or that they act in a redundant fashion in the plant.

Henrik Johannesson, Department of Physiological Botany, Evolutionary Biology Centre, Uppsala University, Villavägen 6, SE-752 36 Uppsala, Sweden

© Henrik Johannesson 2000

ISSN 1104-232X
ISBN 91-554-4816-X

Printed in Sweden by University Printers, Uppsala, 2000
This thesis is based on the following papers, which will be referred to in the text by their Roman numerals:


Reprints of papers I and II were made with permissions from Kluwer Academic Publishers.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABBREVIATIONS</td>
<td>6</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>7</td>
</tr>
<tr>
<td>Arabidopsis thaliana: the model plant</td>
<td>10</td>
</tr>
<tr>
<td>Transcriptional regulation</td>
<td>11</td>
</tr>
<tr>
<td>Homeodomain proteins</td>
<td>12</td>
</tr>
<tr>
<td>Homeodomain proteins in plants: KNOTTED-1 proteins</td>
<td>14</td>
</tr>
<tr>
<td>PHD finger proteins</td>
<td>16</td>
</tr>
<tr>
<td>HDZip proteins</td>
<td>17</td>
</tr>
<tr>
<td>DNA-binding properties of HDZip proteins</td>
<td>23</td>
</tr>
<tr>
<td>Dimerization properties of HDZip proteins</td>
<td>26</td>
</tr>
<tr>
<td>Function of HDZip proteins</td>
<td>28</td>
</tr>
<tr>
<td>RESULTS AND DISCUSSION</td>
<td>19</td>
</tr>
<tr>
<td>Identification of new HDZip genes</td>
<td>23</td>
</tr>
<tr>
<td>Identification of HDZip I target sequences in vitro</td>
<td>24</td>
</tr>
<tr>
<td>ATHB7 and 12 have alternative binding-site specificities</td>
<td>27</td>
</tr>
<tr>
<td>Heterodimer formation between HDZip I proteins</td>
<td>28</td>
</tr>
<tr>
<td>Expression analysis of ATHB5</td>
<td>30</td>
</tr>
<tr>
<td>A reverse genetics approach to understand the function of HDZip proteins</td>
<td>31</td>
</tr>
<tr>
<td>Overexpression of ATHB5 in transgenic Arabidopsis</td>
<td>32</td>
</tr>
<tr>
<td>High-level expression of ATHB13 in transgenic Arabidopsis</td>
<td>33</td>
</tr>
<tr>
<td>Interpretation of phenotypes resulting from constitutive expression</td>
<td>35</td>
</tr>
<tr>
<td>Isolation of T-DNA insertion lines in HDZip I genes</td>
<td>36</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>35</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>36</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>37</td>
</tr>
</tbody>
</table>
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA</td>
<td>abscisic acid</td>
</tr>
<tr>
<td>abi</td>
<td>ABA insensitive mutant</td>
</tr>
<tr>
<td>ACC</td>
<td>1-aminocyclopropane-1-carboxylic acid</td>
</tr>
<tr>
<td>ATHB</td>
<td><em>Arabidopsis thaliana</em> homeobox</td>
</tr>
<tr>
<td>BAC</td>
<td>bacterial artificial chromosome</td>
</tr>
<tr>
<td>bZip</td>
<td>basic leucine zipper</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>HD</td>
<td>homeodomain</td>
</tr>
<tr>
<td>HDZip</td>
<td>homeodomain leucine zipper</td>
</tr>
<tr>
<td>HTH</td>
<td>helix-turn-helix</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance spectroscopy</td>
</tr>
<tr>
<td>SAM</td>
<td>shoot apical meristem</td>
</tr>
<tr>
<td>T-DNA</td>
<td>transferred DNA</td>
</tr>
</tbody>
</table>
INTRODUCTION

Arabidopsis thaliana: the model plant

*Arabidopsis thaliana*, thale cress, belongs to the crucifer family (Brassicaceae). This plant has a broad natural distribution throughout Europe, Asia and North America. In the wild, *Arabidopsis thaliana* grows as a winter annual, that is, the seed germinates in the autumn, pass the winter as a rosette and then flower early during the spring when temperature and light conditions are favourable.

Historically, agriculturally important species such as maize, tomato, potato or barley were the plants of choice for experimental studies. As the efforts of the research community was spread out on several plant species, each considered as the ideal model organism by the different researchers, the progress in the understanding of fundamental plant processes was rather slow (Meinke et al., 1998).

Already in 1943, Friedrich Laibach, Professor at the University of Frankfurt am Main, published an article describing the advantages of *Arabidopsis thaliana* as a research tool. In his article, a number of favourable features of *Arabidopsis* were pointed out, and since then, only a few additional advantages are acknowledged. The features of this plant that make it suitable for experimental studies include several traits. One is the short generation time of *Arabidopsis*; when grown under optimal conditions in the greenhouse it has a generation time of about six weeks. Another advantage is the small size of the plant, which make it possible to grow a large number of plants in limited growth facilities. The diameter of the rosette ranges from 2 to 10 cm in diameter depending on growth conditions, the final height of the inflorescence stem reaches 20-40 and the flowers are about 2 mm long. Furthermore, *Arabidopsis* flowers self-pollinate as the buds open and each plant can produce hundreds of siliques, containing more than 5000 seeds. It is easy to perform genetic crosses of *Arabidopsis* and recover a fully fertile progeny, and both induced and spontaneous mutations can be isolated with ease. Although all these advantages of using *Arabidopsis* for experimental studies were pointed out already in 1943, it would last another forty years before it really became favoured as a model organism (Redei, 1992; Meinke et al., 1998).
Figure 1. Arabidopsis thaliana, thale cress, approximately four weeks old and grown under long day conditions. (Photograph by Eva Söderman).
In 1983, Maarten Koornneef presented a detailed genetic map of *Arabidopsis*. Furthermore, it was demonstrated that the genome of *Arabidopsis* is one of the smallest genomes known among angiosperm species (Meyerowitz and Pruitt, 1985); a characteristic that facilitates the isolation of genes through screening of genomic libraries, which can be a time-consuming task when working with large genomes. Another groundbreaking discovery was that it is possible to transform *Arabidopsis* with *Agrobacterium*, making it possible to generate stable transgenic plants with ease. The methods for *Agrobacterium*-mediated transformation of *Arabidopsis* have developed from tissue culture techniques (Lloyd *et al*., 1986), through seed transformation (Feldmann and Marks, 1987), to the whole-plant transformation methods (Bechtold *et al*., 1993). With the latter method, it is possible to generate large number of transformants with a reasonable effort.

Over the past 20 years, several thousand mutants of *Arabidopsis*, defective in almost every aspect of plant growth and development, have been identified. The development of efficient transformation methods for *Arabidopsis* has facilitated the establishment of large mutant collections of T-DNA insertion lines, carrying randomly inserted T-DNAs. Seed stocks and DNA-pools from these collections of lines are available from public stock centers. The use of these lines for mutant screenings is particularly advantageous as the genes that are disrupted in these lines are tagged by the inserted T-DNA, which facilitate the isolation of the mutated gene.

Research using *Arabidopsis* as a model plant has provided insight into almost every aspect of plant biology. For example, the understanding of phytohormone action has improved drastically through genetic and molecular analysis of *Arabidopsis*. The classical approach to understand the function of plant hormones has largely been based on application of the hormone and recording the observed effect. However, this approach has numerous limitations related to the uptake, transport and sequestration of the exogenously applied hormone. Usually, an indirect, rather than a direct, link between the hormone and its effect on a particular developmental process of the plant will be the result. In contrast, the use of mutants, defective in either their ability to produce the hormone, or in their ability to perceive and respond to the hormone, has been particularly informative (Klee and Estelle, 1991). For example, mutant analysis of the
response to the gaseous plant hormone ethylene lead to the identification of the first hormone receptor in plants (Chang et al., 1993). In addition, increasing knowledge about the signal transduction and physiological responses to a range of different plant hormones has emerged through the genetic analysis of hormone mutants in Arabidopsis (reviewed in McCourt, 1999). The use of Arabidopsis for this purpose makes it possible to isolate and characterize the mutated genes with a reasonable effort due to the extensive amount of genetic information available for Arabidopsis.

The perhaps most important progress of Arabidopsis research is the sequencing of the entire genome, which is performed as a collaborative effort from several participants all over the world. This project was initiated in 1996 and estimated to be completed in year 2004. However, this project has proceeded at a much higher speed than expected; the completion of the genome is now scheduled to the fall, 2000. With this information available, the field of Arabidopsis research will definitively change, making it possible to analyze plants from the perspective of the whole genome.

**Transcriptional regulation**

The process of gene expression links the genetic information that resides in the DNA of the nucleus with the activity of proteins that give the individual cell its characteristics. The conversion of the genetic information into a protein product is to a large extent dependent on the intermediate step, gene transcription. This process is the key-point for the regulation of gene expression, controlling the cell type or timing-specific expression of a certain gene, as well as expression in response to a certain signal (Latchman, 1998). Transcription is regulated by the activity of transcription factors, which interact with target genes and either activate or repress transcription of the gene. In most cases, transcription factors are sequence-specific DNA-binding proteins that bind to specific DNA target sequences and interact with the basal transcription machinery.
**Homeodomain proteins**

Among transcription factors, several different types of DNA-binding protein motifs have been identified and characterized. Homeodomain proteins, encoded by the homeobox gene family, represent one class of transcription factors. Homeobox genes have been shown to be important for the genetic control of development in range of different eucaryotes, i.e. specification of the body plan, determination of cell fate and a number of other developmental processes (for reviews, see e.g. Gehring, 1992; Lawrence and Morata, 1994).

The first homeodomain protein to undergo a structural analysis was the *Drosophila* homeodomain protein Antennapedia (AntP), whose three-dimensional structure in solution was determined by nuclear magnetic resonance (NMR) spectroscopy (Qian et al., 1989; Billeter et al., 1990). The analysis revealed that the homeodomain of AntP is composed of three α-helices that are folded in a tight globular structure. Helix 1 is separated from helix 2 by a loop whereas helix 2 and 3 are separated by a turn. Sequence similarity between the homeodomain and the helix-turn-helix (HTH) motif of prokaryotic regulatory proteins (Pabo and Sauer, 1984), in which the second helix of the HTH motif makes specific contacts with DNA, suggested that homeodomain proteins act as DNA-binding regulatory proteins. Subsequently, X-ray crystallographic studies of homeodomain proteins complexed with DNA have revealed how homeodomain proteins recognize DNA (reviewed in Gehring et al., 1994). Helix 3 are positioned in the major groove of the DNA and make sequence specific interactions with the bases of the DNA. Amino acid residues of the N-terminal extension of the homeodomain perform additional base-contacts in the minor groove of DNA, and amino acid residues positioned in the loop between helix 1 and 2 and at the start of helix 2 make contacts with the DNA backbone. Structural studies show that various homeodomain proteins from a range of different organisms interact with DNA in essentially identical ways, demonstrating that the mode of DNA binding of homeodomain proteins has been highly conserved during evolution.
**Homeodomain proteins in plants: KNOTTED-1 proteins**

The first homeodomain-encoding gene to be isolated from plants was *KNOTTED-1* from maize (Vollbrecht *et al*., 1991) corresponding to the *Kn1* locus, which is defined by several dominant gain-of-function mutations that disrupt leaf development, resulting in the formation of knots on the leaf blade. In the *Kn1* mutants, *KNOTTED-1* is expressed ectopically in vascular bundles within developing leaves, thus correlating with the phenotypic alterations evident in the mutant (Smith *et al*., 1992). The wild-type *KNOTTED-1* gene is expressed in the apical meristem of vegetative and floral shoots but expression is excluded from determinate organs such as leaves (Smith *et al*., 1992), suggesting that this gene is important for shoot meristem function. Mutations in the *Arabidopsis KNOTTED*-like gene *SHOOTMERISTEMLESS (STM)* result in seedlings that are incapable of developing shoot apical meristems during embryogenesis (Barton and Poethig, 1993; Long *et al*., 1996). These observations suggest that *KNOTTED*-like genes have important roles in shoot apical meristem maintenance (Long *et al*., 1996).

**PHD finger proteins**

A different class of plant homeodomain proteins are defined by virtue of a second conserved protein motif denoted PHD-finger. An example of a PHD-finger homeodomain protein is HAT3.1 from *Arabidopsis thaliana*. This protein was isolated based on its capability to interact with the light-induced *cab-E* promoter of tobacco (Schindler *et al*., 1993). The PHD-finger class of homeodomain proteins also includes the maize protein Zmhox1a, which binds to the 26 bp feedback control element of the of the *Shrunken* gene promoter (Bellman and Werr, 1992), and PRHA/PRHP from *Arabidopsis* and parsley, respectively, that both interact with a region of the *pr2* promoter required for elicitor-mediated expression of this gene (Korfhage *et al*., 1994).

**HDZip proteins**

In an effort to isolate homeobox containing genes from *Arabidopsis thaliana*, a strategy previously used to isolate homeobox genes from the nematode *Caenorhabditis elegans*
(Bürglin et al., 1989) was applied. *Arabidopsis* cDNA libraries were screened with degenerate oligonucleotide probes, designed to match all possible codon combinations corresponding to eight conserved amino acids in the helix 3 region of the homeodomain (Ruberti et al., 1991; Mattsson et al., 1992; Schena and Davies, 1992). A number of clones were recovered and showed extensive sequence similarity to known homeobox genes and were designated *Arabidopsis thaliana* homeobox genes (ATHB). In addition, these clones contained a second sequence element positioned 3’ and closely linked to the homeobox, encoding an α-helical stretch of amino acids, with a periodic repetition of leucine residues at every seventh position. This configuration of aminoacid residues resembled that of the transcription factor C/EBP, in which the leucine residues are arranged along one side of the helix and interact with the leucine residues projecting from a second helix, forming a dimeric “leucine zipper” complex (Landschulz et al., 1988). As the positions of putative DNA-binding units relative the leucine zipper domain are similar in ATHB and bZIP proteins, respectively, it was proposed that proteins with a contiguous homeodomain-leucine zipper architecture should be referred to as HDZip proteins (Ruberti et al., 1991).

Additional members of the HDZip gene family have subsequently been isolated from *Arabidopsis* (Carabelli et al., 1993; Söderman et al., 1994; Schena and Davis, 1994; Baima et al., 1995; Lu et al., 1996; Di Cristina et al., 1996; Sessa et al., 1998; Lee and Chun, 1998; Kubo et al., 1999; Zhong and Ye, 1999; Hanson (II); Johannesson et al., (IV). HDZip genes have also been identified from a range of other plant species; tomato (Tornero et al., 1996; Meissner and Theres, 1995; Mayda et al., 1999, sunflower (Chan and Gonzalez, 1994; Gonzalez et al., 1997), carrot (Kawahara et al., 1995; Mattsson et al., 1995), soybean (Moon et al., 1996), *Pimpinella brachycarpa* (Moon et al., 1996), resurrection plant (Frank et al., 1998) and ferns (Aso et al., 1999). HDZip genes have been isolated from plants, but not from any other eukaryotes, suggesting that the HDZip class of homeodomain proteins is specific to plants.
**DNA-binding properties of HDZip proteins**

Based on sequence criteria and supported by intron positions, HDZip proteins have been grouped in four different families, HDZip I – IV (Sessa *et al.*, 1994). Proteins that belong to HDZip I and II are very similar to each other with respect to the architecture of the HDZip domain. DNA-binding studies of ATHB1 (HDZip I) and 2 (HDZip II)

**Figure 2.** A hypothetical model of a HDZip protein binding to DNA, based on the three-dimensional structures of the *Drosophila* engrailed homeodomain and the yeast GCN4 leucine zipper motif, respectively (K. Johansson, H. Johannesson and E. Söderaman, unpublished).
revealed that these two proteins bind as homodimers to pseudopalindromic binding sites, CAAT(A/T)ATTG and CAAT(G/C)ATTG, respectively, consisting of two 5 bp half-sites that overlap at the central position (Sessa et al., 1993). Furthermore, it was obvious that the spacing between the homeodomain and the leucine zipper is critical for the function of HDZip proteins as an insertion of two amino acids in this region of the HDZip protein abolished its DNA-binding activity. Two HDZip II proteins from other species other than Arabidopsis thaliana, Oshox1 from rice and CPHB-1 from the resurrection plant Craterostigma plantagineum, bind DNA with specificities similar to that of ATHB2 (Meijer et al., 1997; Frank et al., 1998).

The class III proteins ATHB8, 9, 14 have a different architecture of their HDZip domains as compared with HDZip I and II proteins, with insertions of four amino acids both between helix 2 and 3 of the homeodomain and between helix 3 and the leucine zipper domain (Baima et al., 1995; Sessa et al., 1998). Thus, the spacing between the homeodomain and leucine zipper domain is different in these proteins as compared to that of HDZip I and II proteins, which suggests that class III proteins might exhibit a different mode of DNA-binding. However, DNA binding studies of the ATHB9 protein revealed that this protein interacts specifically with the DNA sequence GTAAT(G/C)ATTAC, thus demonstrating that the HDZip domain of class III proteins interacts with DNA in a similar fashion as that of class I and II protein.

ATHB10/GLABRA2, a class IV protein, is identical to class I and II proteins in the architecture of the homeodomain and spacing between the homeodomain and leucine zipper domain (Di Christina et al., 1996). However, the leucine zipper motif of ATHB10/GLABRA2 is separated in two subdomains by a loop of 10 amino acid residues. Replacement of the ATHB2 leucine zipper domain with that of ATHB10/GLABRA2 resulted in a chimeric protein that was capable of interacting with the ATHB2 recognition site. Hence, despite the different organization of the leucine zipper domain of ATHB10/GLABRA2, this protein is active in dimerization and most likely exhibits a similar mode of DNA-binding as class I and II proteins.

The DNA-binding activities in vivo have been analyzed for three HDZip proteins: ATHB1 and 2 from Arabidopsis (Aoyama et al., 1995; Steindler et al., 1999) and Oshox1 from rice (Meijer et al., 1997). ATHB1 was able to activate transcription
from a promoter containing the ATHB1 target sequence CAAT(A/T)ATTG upstream of a reporter gene in tobacco cells. On the other hand, Oshox1 and ATHB2 were shown to repress reporter gene activity in rice suspension cells and *Arabidopsis* leaves, respectively. The transcriptional repression property of Oshox1 was lost upon deletion of the N-terminal part of the protein, indicating that the repression function of Oshox1 resides in this part of the protein (Meijer et al., 1997). Thus, based on these observations it is obvious that the HDZip family of transcription factors include both activators and repressors of transcription.

In another study, it was shown that the N-terminal part of the protein, which repress transcription in rice cells, functions as a transcriptional activator in yeast cells, suggesting that Oshox1 might function as either a transcriptional activator or repressor depending on the promoter context of the target gene (Meijer et al., 2000). To what extent this dual function applies to other HDZip proteins awaits further testing.

**Dimerization properties of HDZip proteins**

Protein-protein interaction studies *in vitro* have shown the ATHB1 and ATHB2 proteins to exclusively form homodimers (Sessa *et al*., 1993). Similarly, studies on the interaction between the sunflower HDZip proteins Hahb-1 and -10, which belong to the class I and II, respectively, demonstrated that these proteins were only capable of homodimerization (Gonzalez *et al*., 1997). Heterodimerization has been demonstrated between two HDZip II proteins from different species; Oshox1 and ATHB2 (Meijer *et al*., 1997) and between the two HDZip II proteins CPHB-1 and -2 from the resurrection plant *Craterostigma plantagineum* (Frank *et al*., 1998). Together, these data suggest that heterodimerisation is possible between members of the same class, but not between proteins of different classes. The mechanism behind this apparent specificity in dimerization is not clear. However, the organization of amino acid residues in the N-terminal part of the leucine zipper is different between HDZip I and II proteins and is probably of great importance (Gonzalez *et al*., 1997).
Function of HDZip proteins

The function of three Arabidopsis HDZip genes has been resolved through mutant analyses. The glabra2 mutation results in defective trichome and root hair development (Rerie et al., 1994; Masucci et al., 1996). The GLABRA2 gene was later shown to be identical to the HDZip IV gene ATHB10 (Di Christina et al., 1996). Mutations in the ANL2 gene (HDZip IV) result in abnormal anthocyanin accumulation in subepidermal tissues of rosette leaves and disturbed organization of primary root cells (Kubo et al., 1999). Disruption of the IFL1 gene (HDZip III) abolishes the formation of normal interfascicular fibers in Arabidopsis inflorescence stems (Zhong et al., 1999).

Functional information on HDZip class II genes is based on results from expression analyses and analyses of transgenic plants with altered expression levels of the genes. Expression of the ATHB2 (HDZip II) gene is induced by changes in the ratio of red to far-red light (Carabelli et al., 1996) and the phenotypic analysis of transgenic Arabidopsis plants with altered levels of ATHB2 suggests this gene to be a mediator of shade avoidance responses (Steindler et al., 1999). Two dehydration-stress inducible HDZip II genes, CPHB-1 and -2, were isolated from the resurrection plant, Craterostigma plantagineum, using a differential display technique. Expression of both genes are induced at early stages of dehydration suggesting that these genes may be involved in the regulation of gene expression during dehydration (Frank et al., 1998). Overexpression of the rice Oshox1 protein (HDZip II) protein in transgenic Arabidopsis resulted in small and lancet-shaped leaves (Meijer et al., 1997). Oshox1 is predominately expressed in leaves and together, these data are suggestive of a role for this gene in the control of leaf morphogenesis.

Functional information on HDZip class I proteins is limited. ATHB6, -7 and -12 are induced by exogenous ABA, water deficit and osmotic stress. The induction of ATHB6 and -7 by drought was strictly dependent on ABA as no increase in ATHB6 and -7 transcripts was detected in the drought-treated seedlings of the ABA-deficient mutant aba-3. Furthermore, the induction by ABA of ATHB7 was impaired exclusively in the ABA-response mutant abil whereas the induction of ATHB6 was impaired in both abil and abi2. These data suggest that ATHB6, 7 and 12 probably act in an ABA-dependent signal transduction pathway, mediating the growth response to drought in the plant.
Transgenic plants with enhanced levels of *ATHB1* are defective in the palisade parenchyma as cells of this tissue in the transgenic plants are replaced by other cells, similar to spongy mesophyll cells. In addition, the transgenic plants exhibited deetiolated phenotypes in the dark as well as abnormal expansion of cotyledons. Together, these observations suggest that *ATHB1* is involved in the activation of target genes that are closely linked to plant development (Aoyama *et al*., 1995). Antisense suppression of the tomato HDZip I gene *H52* in transgenic tomato plants results in necrosis and abscission of leaves (Mayda *et al*., 1999). The accumulation of ethylene and salicylic acid is increased and several defense-related genes are constitutively expressed in the transgenic plants. *H52* gene expression is induced upon infection with virulent pathogens. Thus, these data are suggestive of a role of *H52* as a transcriptional regulator of genes involved in programmed cell death (PCD), acting to protect the cells against PCD (Mayda *et al*., 1999).
RESULTS AND DISCUSSION

Identification of new HDZip genes (II, IV)

By searching available databases, we were able to identify 26 *Arabidopsis* sequences that show similarity to HDZip class I and II genes. Of these, 17 corresponded to previously characterized HDZip genes (ATHB1, 2, Ruberti et al., 1991; ATHB3, Mattson et al., 1992; ATHB5, 6, 7, Söderman et al., 1994; ATHB12, Lee and Chun, 1998; ATHB13, Hanson et al., paper III; ATHB16, Y. Wang, in preparation; ATHB4, Carabelli et al., 1993; HAT22, Schena and Davies, 1992; HAT1, 2, 3, 9, 14, Schena and Davies, 1994, ATHB17, Ruberti et al., unpublished), whereas nine represented novel sequences encoding putative HDZip proteins. These nine genes were designated ATHB20, 21, 22, 23, 40, 51, 52, 53 and 54, respectively.

Biochemical characterization of HD-DNA complexes have shown a majority of HD proteins to recognize TAAT-containing DNA-sequences (reviewed in Treisman et al., 1992). The inherent similarities in DNA-binding specificity of all these HD proteins implicate a structural conservation at the protein-DNA interface, which has been confirmed by structural studies of several HD-DNA complexes (reviewed in e.g. Gehring et al., 1994). A homeodomain consensus sequence has been defined based on a compilation of 346 homeodomain sequences (Bürglin, 1994). The amino acids conserved in this consensus sequence can be assigned to three different functions: contributing to the hydrophobic core, to sugar-phosphate backbone recognition or specific interactions with the bases in the TAAT core site. The HDs of all HDZip proteins, including ATHB20, 21, 22, 23, 40, 51, 52, 53 and 54, contain the five invariant amino acid (L_{16}, W_{48}, F_{49}, N_{51}, R_{53}) and six out of seven of the highly conserved residues (F_{20}, L_{26}, L_{40}, V_{45}, I_{47}, L_{57}) of this homeodomain consensus sequence. In addition to the conserved homeodomain, these protein, including ATHB20, 21, 22, 23, 40, 51, 52, 53 and 54, contain leucine zipper motifs in identical positions C-terminal to the homeodomain.

An initial alignment of the HDZip region of the deduced aminoacid sequences revealed that the ATHB22 sequence did not align to the other HDZip sequences in the region upstream of helix 2 (Figure 3). However, if 8 aminoacids corresponding to the region between helix 1 and 2 were looped out, a perfect alignment of ATHB22 along the
Figure 3. Alignment of the amino acid sequences of HDZip proteins. Gaps are indicated by dashes.

A homodomain consensus sequence based on a compilation of 346 homeodomain sequences (Bürglin, 1994) is indicated below the alignment. The three α-helices of the homeodomain (Bürglin, 1994) are shown below the alignment.
entire HDZip domain was possible. A similar atypical architecture of the homeodomain is evident for the yeast homeodomain MATα2, which has an insertion of three aminoacids in the turn between helices 1 and 2 (Hall and Johnson, 1987). The structure of MATα2 (Wolberger et al., 1991) is very similar to the Drosophila homeodomain proteins engrailed (en) and AntP structures, indicating that the extra aminoacids in the MATα2 homeodomain do not influence the overall structure of this domain. Thus, despite the atypical architecture of the ATHB22 homeodomain, it is likely that this protein adopts a similar structure and exhibits a similar DNA-binding property as those of other HDZip proteins.

Based on sequence criteria, HDZip proteins have been grouped in four distinct subfamilies, HDZip I – IV (Sessa et al., 1994) and the members of each subfamily are also related with respect to intron positions. A number of positions separate HDZip I and II proteins. More specifically, position 46 of helix 3 is invariant within each of the classes but distinct between HDZip I and II. The same is true for position 58 in the region between the homeodomain and the leucine zipper domain. The alignment shows that, according to these criteria, the nine newly identified proteins belong to HDZip I, since the amino acid residues in positions 46 and 58 are identical to those of all the previously identified HDZip I proteins.

The evolutionary relationship between the entire set of HDZip I and II aminoacid sequences was analyzed by parsimony, using PAUP (Swofford, 2000). The phylogenetic analysis resulted in four equally parsimonious trees. Within the HDZip domain, introns can be found in five different positions. The intron positions are conserved within the different subclasses defined by the tree and thus support the major branching pattern of the tree. The tree revealed two distinct classes, corresponding to the HDZip I and II classes previously identified by Sessa et al., (1994), respectively. All of the nine newly identified sequences were included in HDZip class I, confirming that ATHB20, 21, 22, 23, 40, 51, 52, 53 and 54 belong to HDZip I.

The chromosomal locations of each of the HDZip genes were determined based on the positions of corresponding BAC-clones on the physical map of Arabidopsis thaliana. As shown in Figure 4, HDZip I and II genes are evenly distributed over the entire length of the five chromosomes of Arabidopsis.
Figure 4. Locations of HDZip I and II genes on the five chromosomes of *Arabidopsis thaliana*. The position of each gene is according to the physical map of *Arabidopsis* and their locations are indicated relative a subset of mi-RFLP markers (Liu et al., 1996)
Identification of HDZip target sequences \textit{in vitro} (I)

HDZip proteins contain leucine zipper domain located immediately C-terminal to the homeodomain. As the leucine zipper motifs in other classes of transcription factors mediate dimerization it was suggested that HDZip proteins are active as protein dimers in the binding of DNA, such that the leucine zipper juxtaposes a pair of DNA interacting units onto the DNA (Ruberti \textit{et al.}, 1991). This notion was confirmed when analyzing the DNA-binding properties of ATHB1 and ATHB2, which interacted with the distinct pseudopalindromic sequences CAAT(A/T)ATTG and CAAT(G/C)ATTG, respectively, as homodimers (Sessa \textit{et al.}, 1993). Thus, ATHB1 and ATHB2 interact with nearly identical DNA sequences and the principal difference is their specificity for the central position in the recognition site. Considering the high degree of sequence similarity within the homeodomain of the HDZip proteins, it is possible that these proteins interact with very similar DNA binding sites. Moreover, it is likely that proteins that belong to the same subclass exhibit identical binding site preferences, such that all HDZip I and II proteins have the same specificity as ATHB1 and ATHB2, respectively. To test this hypothesis, the binding specificity for a number HDZip class I proteins was analyzed, starting with ATHB5. A PCR assisted binding-site selection procedure was used to select for high-affinity binding sites for ATHB5 from a pool of DNA-fragments containing a 12 bp random core sequence. Strikingly, the majority of the clones contained the pseudopalindromic sequences CAAT(A/T)ATTG or CAAT(G/C)ATTG, confirming that the HDZip I protein ATHB5 binds DNA in a similar fashion as ATHB1. In contrast to ATHB1, ATHB5 did not discriminate between binding sites in which the identity of the central positions was different as the frequency of recovered clones containing either an A/T or a G/C pair in the central position was essentially the same. In electrophoretic mobility shift assays, ATHB6 and 16 interacted with similar affinity to either A/T or G/C containing sequences. This result demonstrates that ATHB6 and 16, similar to ATHB5, are independent on the identity of the central position in the binding site. In contrast, ATHB1 showed a distinct and ATHB3 and 13 exhibited a slight preference for an A/T pair in the central position specificity. Thus, it is evident that HDZip proteins interact with very similar binding sites, but central position specificity appear to vary among HDZip class I proteins.
What is the determinant of central position specificity in the HDZip proteins? A DNA binding model of HDZip proteins presented by Sessa et al., (1993) predicts that HDZip and bZip proteins interact similarly with DNA and that Arg55 of the HDZip domain is functionally equivalent to Arg243 of the yeast bZip protein GCN4. When complexed with the GCN4 recognition site, TGA(C/G)TCA, Arg243 of one GCN4 monomer donates hydrogen bonds to the guanidine of the central pair, whereas Arg243 of the other monomer makes phosphate backbone interactions (Ellenberger et al., 1992).

The importance of Arg55 for central position specificity of HDZip proteins was tested experimentally by a mutational analysis of the ATHB2 HDZip domain (Sessa et al., 1997). The results indicated that Arg55 of ATHB2 is important for central position specificity, but residues outside helix 3 also contribute to binding-specificity. Thus, specificity for the central position may be attributable to the conformation of the Arg55 side chains in the HDZip domain. This finding is consistent with the fact that Arg55 is invariant among all HDZip proteins and therefore, this position alone can thus not explain the apparent diversity of central position preference for HDZip proteins.

**ATHB7 and 12 have alternative binding site specificities (I)**

In contrast to ATHB1, 3, 5, 6, 13 and 16, neither ATHB7 nor ATHB12 interacted with the HDZip consensus site CAATNATTG. This was a surprising result considering the fact that all HDZip I and II proteins, including ATHB7 and 12, are highly similar in sequence within the homeodomain and it is rather unlikely that these proteins would differ dramatically in their DNA binding specificities. One possible explanation for the result is that the ability of the ATHB7 and 12 proteins to bind DNA is dependent on correct post-translational modifications, e.g. protein phosphorylation. This possibility is consistent with the presence of a potential phosphorylation site at a position in the homeodomains of ATHB7 and 12, which is absent in homeodomains of other class 1 HDZip proteins. Furthermore, the transcription of both genes are induced by ABA, osmotic stress and drought conditions (Söderman et al., 1996; Lee and Chun, 1998), and in the case of *ATHB7* by a signaling mechanism that involves protein phosphorylation/
dephosphorylation (Söderman et al., 1996), suggesting that this signaling pathway might also act on ATHB7 directly on the protein level by a similar mechanism.

Another possibility is that ATHB7 and 12 interact with different binding-sites than those of other HDZip I proteins. In support of this hypothesis, ATHB12 has recently been demonstrated to specifically interact with the palindromic sequence TCAATTAATTGA (Chun and Lee, 1998, US Patent No 5,981,729). A closer inspection of the ATHB12 recognition site reveals that it consists of two non-overlapping TCAATT half-sites, as opposed to the HDZip consensus-binding site CAATNATTG, which consists of two CAATN half-sites overlapping in the central position. Thus, the principal difference between the two sites is the spacing between the half-sites, whereas the half-sites as such are essentially identical. This finding suggests that the interaction of ATHB12 at the protein-DNA interface is the same as for other HDZip proteins, but that ATHB12 adopts a slightly different structure that affects the spacing between half-sites. Furthermore, considering the high degree of sequence similarity between ATHB7 and 12, it is likely that ATHB7 exhibits an identical DNA-binding specificity as that of ATHB12.

The mechanism by which half-site spacing in HDZip proteins is determined is not clear. In bZip proteins, it has been shown that the short surface spanning the fork region, which connect the leucine zipper and the DNA-binding basic region, is the major determinant of half-site spacing (Kim and Struhl, 1995). However, the region of HDZip proteins that corresponds to the fork region in bZip proteins is conserved among all HDZip proteins, including ATHB7 and 12, suggesting that this mechanism of half-site spacing determination is not applicable to HDZip proteins.

Another possibility is that HDZip proteins, in analogy with e.g. the *Drosophila* homeodomain proteins AntP and en, perform DNA contacts in the minor groove of DNA through arginine residues at positions 3 and 5 in their N-terminal arms of the homeodomain (reviewed in Gehring et al., 1994). A closer inspection of the region of the homeodomain preceding helix 1, all HDZip class proteins, except ATHB7 and 12, have a basic stretch of amino acids, which shows resemblance to the N-terminal extension of classical homeodomain proteins. Although not tested experimentally, it is possible that the N-terminal region of HDZip proteins is functionally equivalent to that of AntP and
en and determines DNA-binding specificity through minor groove interactions. If that were the case, the alternative spacing preference of ATHB7 and 12 could be a consequence of the alternative architecture of the N-terminal region.

In summary, HDZip proteins are similar in sequence within their HDZip domain, which suggest that they should interact with very similar DNA sequences. Nevertheless, the diversity in central position specificity among HDZip proteins, manifested either as preference for a certain pair in this position or as preference for an alternative spacing between half-sites, appear to be larger than initially expected. Thus, the functional differences between different HDZip proteins may in part be due to this diversity in their DNA-binding.

Heterodimer formation between HDZip class I proteins (I).

The fact that some of the HDZip proteins have very similar DNA-binding specificities and yet appear to have distinct biological functions raise the hypothesis that the targeting of HDZip proteins in vivo to the correct promoters is dependent on interactions with other proteins. A special case of such protein-protein interaction is the formation of heterodimers between different members of HDZip proteins.

The interaction between a recombinant bacterial ATHB5 protein and in vitro translated 35S-labeled HDZip proteins was analyzed and the result showed that ATHB5 interacted with ATHB6, 7, 12 and 16 with different apparent affinity, but not with ATHB1, suggesting that HDZip proteins show selectivity in dimer formation, specified by the dimerization domain.

In bZip proteins, such as the protein products of the proto-oncogenes c-fos and c-jun, the leucine zipper domain itself is capable of dictating specificity (O’Shea et al., 1992). Specificity depends on electrostatic interactions between amino acids of e and g positions of the heptad repeats (abcdefg)n that form the hydrophobic interface in the two monomers, which is consistent with the known crystal structures of GCN4 (Ellenberger et al., 1992) and a Fos/Jun heterodimer (Glover and Harrison, 1995). An interhelical salt bridge hypothesis (Vinson et al., 1993), based on functional analyses of leucine zippers suggests that the identity of amino acid in the g and e position of each
unit determine which proteins that have the potential to dimerize. Applied to the class I HDZip proteins, this hypothesis predicts the proteins to have the capacity to form both homo- and heterodimers. The in vitro data on ATHB5 indicate that the protein in fact does show selectivity in dimer formation, suggesting that HDZip proteins rely on a mechanism different from that described for bZip proteins to dictate dimerization specificity.

The selective heterodimer formation between HDZip class I proteins might provide the plant with transcription factors that, even when interacting with the same target sequence as the original homodimer, have distinct activation characteristics. Thus, it is possible that HDZip proteins constitute a functional complex of homo- and heterodimeric transcription factors in the plant, providing a mechanism by which the plant can increase the level of complexity in transcriptional regulation on the protein level.

Expression analysis of ATHB5 (II)

A first step to understand the function of HDZip proteins is to determine the tissue and timing specificity of HDZip gene expression. In the case of ATHB5, this gene was previously shown to be expressed at moderate levels in adult leaf, stem and root tissues (Söderman et al., 1994); however, the expression of ATHB5 at other developmental stages was not analyzed. Analysis of the expression of ATHB5 during early development showed that ATHB5 gene transcription was turned on early after the onset of germination and then gradually declined with increasing age of the seedlings. Additional experiments revealed that ATHB5 mRNA was detectable as early as 12 h post-imbibition, but not in the dry seed (data not shown). Histochemical staining of transgenic Arabidopsis seedlings harboring an ATHB5::GUS fusion revealed strong GUS activity specifically in the hypocotyl of germinating seedlings, concentrated around the transition zone, which marks the boundary between the hypocotyl and the root. Staining was retained at the transition zone as the seedlings grew older, but was also visible with low intensity in the petioles of cotyledons and developing true leaves. Weak GUS staining associated to vascular tissue was found in root, stem and cauleine leaf
tissues and no visible staining was found in flower organs. Thus, the expression analysis of \textit{ATHB5} showed that the activity of this gene appears to be under strict developmental control and activated by the initiation of germination.

The staining intensity at the transition zone of the \textit{ATHB5}::GUS transgene decreased markedly when seedlings were treated with 10 µM ABA for 12 hrs. This effect was specific for ABA as the staining intensity was unaffected by the treatment with auxin, gibberellic acid, cytokinin or the ethylene precursor ACC. These data indicate that \textit{ATHB5} gene expression is dependent on ABA signaling. This notion was supported by results from experiments assessing whether \textit{ATHB5} expression during postgerminative growth was dependent on the activity of known ABA response loci. The expression level of \textit{ATHB5} was markedly lower in \textit{abi3-1} and \textit{abi5-1} mutant background as compared to the wild type, demonstrating that \textit{ATHB5} represents a target gene for transcriptional regulation by \textit{ABI3} and \textit{ABI5} during postgerminative growth, either directly or indirectly.

In summary, the expression analysis allows the distinction of cell types and developmental stages where the gene primarily exerts its function. In the case of \textit{ATHB5}, we can conclude that this protein is functional primarily in cells of the hypocotyl during seed germination. Moreover, the fact that the \textit{ATHB5} expression level is affected by ABA and lower in the ABA response mutants \textit{abi3} and \textit{abi5} indicate that the function of \textit{ATHB5} is related to ABA signaling.

**A reverse genetics approach to understand the function of HDZip proteins**

Approaches such as gene expression analysis to understand gene function are only correlative and do not necessarily provide a direct link between a gene and its function. In contrast, by the analysis of null mutations in the genes of interest it is possible to establish their function simply by assessing the phenotypic effects resulting from the mutation. This approach has had an enormous impact on the understanding of developmental processes in a range of different organisms (e.g. Nüsslein-Volhard, 1994; Burns \textit{et al.}, 1994; Spradling \textit{et al.}, 1995; Brandon \textit{et al.}, 1995).
Mutations created by insertions of either T-DNA or transposable elements are particularly advantageous as they provide direct means by which the mutated gene can be isolated. Individual lines bearing insertions in the gene of interest are easily identified by amplification of flanking regions of the inserted element using polymerase chain reaction (PCR) with a combination of gene-specific primers and primers complementary to border sequences of the inserted element. This method was originally developed for Drosophila where PCR was used to screen for P-element insertions (Ballinger and Benzer, 1989; Kaiser and Goodwin, 1990) and has also been utilized in other organisms such as Caenorhabditis elegans (Zwaal et al., 1993). In Arabidopsis, this approach has been used to isolate plants carrying insertions in a range of different genes (e.g. McKinney et al., 1995; Krysan et al., 1996; Winkler et al., 1998). Considering the relatively small genome size of Arabidopsis, it has been estimated that it is sufficient to screen around 100,000 lines in order to find a plant carrying an insert in any gene of interest (Azpiroz-Leehan and Feldmann, 1997).

A limitation to screens for loss-of-function mutations is that the recovered mutants often lack discernable phenotypes, due to genetic redundancy. Redundancy is evident among members of the ethylene receptor gene family (Hua and Meyerowitz, 1998) as well as among homeotic genes that control flower development in Arabidopsis thaliana (reviewed in Martienssen and Irish, 1999). An alternative strategy to reveal gene function is to assess the phenotypic effect of enhanced or reduced levels of gene activity by expressing the gene in either sense- or antisense orientation from the strong constitutive promoter of the cauliflower mosaic virus (CaMV) 35S gene. The anti-sense approach will suffer from the same problem as that of insertion lines, namely gene redundancy, unless the expression of the anti-sense construct has the capacity to inactivate several genes simultaneously that are related in sequence and act in a redundant fashion. In contrast, overexpression of the gene can provide functional information of a gene, which will be manifested through a visible phenotype. One complication of high-level expression from the 35S-promoter is that it results in gene activity in tissues or developmental stages where the gene is not normally expressed and hence, resulting in phenotypic deviations from wild type that are not related to the
function of the gene. However, a careful expression analysis of the gene should facilitate the distinction of phenotypes that reflect the actual function of the gene.

**Overexpression of ATHB5 in transgenic Arabidopsis (III)**

Phenotypic analysis of ATHB5 overexpressor lines (35S::ATHB5) revealed that these plants did not differ from wild type in their overall growth and development. However, the 35S::ATHB5 plants differed from wild type in their seedling morphology. Hypocotyls of young seedlings were shorter and radially expanded at the base. In addition, the primary roots as well as petioles of cotyledons of 35S::ATHB5 seedlings were shorter than those of the wild type. Thus, the phenotypic deviations from wild type seen in the 35S::ATHB5 plants are consistent with the results from the expression analysis of ATHB5, which suggest that it is likely that the ATHB5 gene exerts its function primarily in hypocotyls of germinating seedlings and not in the adult plant. The effect of the plant hormone ABA on the intensity of GUS staining in the ATHB5 promoter GUS-fusion plants, as well as the downregulation of ATHB5 in the ABA response mutants abi3 and abi5, suggested that overexpression of ATHB5 in transgenic plants might affect the response of the seedlings to growth inhibition by ABA. This notion was confirmed by experiments assessing the sensitivity of wild type and 35S::ATHB5 plants to the inhibitory effects of exogenously applied ABA on seed germination and seedling root growth. The results from these assays showed that the 35S::ATHB5 line was more sensitive to ABA at all concentrations tested as compared with the wild type. The ABA concentration needed to achieve 50% inhibition of either seed germination or root growth was two-fold lower for the 35S::ATHB5 line than for the wild type. These results suggest that overexpression of ATHB5 interferes with ABA signaling, resulting in seedlings hypersensitive to ABA. The finding that the 35S::ATHB5 seedlings were hypersensitive to ABA was consistent with the results from experiments assessing ABA regulated gene expression, demonstrating that the ABA responsive gene RAB18 was hyperinduced by ABA in 35S::ATHB5 seedlings as compared to wild-type.
During seed development the level of endogenous ABA peaks during mid to late embryogenesis before returning to low levels in the dry seed (Rock and Quatrano, 1995). ABA has been demonstrated to be important for the establishment of seed dormancy since *Arabidopsis* mutants either defective in ABA biosynthesis or in the ability to sense ABA fail to become dormant (Koorneef *et al*., 1982; 1984). Furthermore, mutations that enhance the sensitivity to ABA confer hyperdormancy (Cutler *et al*., 1996; Gosti *et al*., 1999). Consequently, seeds from 35S::*ATHB5* plants that are hypersensitive to exogenous ABA should therefore also be hyperdormant. However, overexpression of *ATHB5* does not affect seed dormancy, indicating that *ATHB5* does not regulate ABA sensitivity during seed development and consequently does not take part in the regulation of seed dormancy. This hypothesis is consistent with the fact that *ATHB5* is not expressed in wild-type plants during seed maturation when dormancy is established.

In summary, the functional analysis of the HDZip gene *ATHB5* provides evidence for that this gene is active in the regulation of germination and postgerminative growth, possibly as an ABA dependent negative regulator of cell expansion. Clearly, germination and postgerminative growth is highly dependent on ABA signaling and it has been suggested that the inhibitory effect of ABA on these processes is related to the inhibition of storage reserve mobilization, since this hormone alters the pools of available nitrogen and energy as well as prevents the degradation of seed storage proteins (Garciarrubio *et al*., 1997). It is possible that the mechanism by which *ATHB5* regulate germination and postgerminative growth is through restriction of energy and metabolites as a component of an ABA response pathway in *Arabidopsis thaliana*.

High-level expression of *ATHB13* in transgenic *Arabidopsis* (II)

Increased expression levels of *ATHB13* in transgenic plants (35S::*ATHB13*) did not result in any major phenotypic deviations in overall growth and development as compared to the wild type. However, when grown in vitro on standard medium, the 35S::*ATHB13* seedlings differed from wild type by decreased cotyledon width. It was evident that this effect was mediated by sucrose as the phenotypic deviation
disappeared when metabolizable sugars were omitted from the growth medium. The effect of sucrose on the cotyledon width of the 35S::ATHB13 seedlings was dose-dependent, since cotyledon width decreased with increasing sucrose concentrations. Measurements of the size of individual epidermal cells demonstrated that the sucrose-dependent decrease in cotyledon width of 35S::ATHB13 seedlings was a consequence of inhibited epidermal cell expansion. The effect of ATHB13 overexpression on cotyledon cell size and shape is most likely dependent on sucrose signaling, since neither non-metabolizable sugars nor genetic crossings with transgenic plants with altered hexokinase activities affected the morphology of 35S::ATHB13 seedlings. This hypothesis is consistent with results from experiments analyzing sugar regulated gene expression, which demonstrated that the expression of a subset of known sugar-responsive genes were induced by sucrose to a higher level in 35S::ATHB13 seedlings than in the wild-type. Taken together, these data show that ATHB13 affect both morphology and gene regulation in response to sucrose, thus representing the first transcription factor shown to directly take part in a sucrose signaling pathway.

**Interpretation of phenotypes resulting from constitutive expression (II, III)**

As stated above, high-level expression from the 35S-promoter can result in ectopic gene activity, i.e. the gene is active in tissues or stages where it is not normally active. In the case of the 35S::ATHB5 plants however, the expression analysis show this gene to be active in the hypocotyl early after the onset of germination and dependent on the ABA response genes ABI3 and ABI5, consistent with the ABA hypersensitivity phenotype during seed germination conferred by the gene when expressed at high levels. In addition, the fact that the phenotypic deviations seen in the ATHB5 overexpressors are specific for the germination stage suggests that the target genes, which ATHB5 activates, are only accessible during this developmental stage. Thus, it is possible that ATHB5 needs to interact with other factors that are active during postgerminative growth in order to be functional. In contrast, the ATHB13 gene is not expressed in cotyledons where it exerts its effect when overexpressed (J. Hanson, in preparation). Thus, the effect of ATHB13 on cotyledon cell size and shape is most likely a consequence of ectopic expression and
does not directly reflect the function of the gene in the wild-type plant. This hypothesis is supported by the observation that \textit{ATHB23}, a close relative to \textit{ATHB13}, is expressed in cotyledons and has a similar effect as \textit{ATHB13} on cotyledon cell shape and size when overexpressed (J. Hanson, in preparation).

\textbf{Isolation of T-DNA insertion lines in HDZip class I genes (II, III, IV)}

In order to obtain loss- or reduction-of-function plants of HDZip genes, the anti-sense approach has been used for a number of HDZip genes. However, this method has proved to be inefficient for a majority of the genes tested. Instead, PCR based screens for insertion mutants in HDZip I genes were performed. A population of 6000 T-DNA insertion lines with an average of 1,5 insertion events per line has been deposited at the \textit{Arabidopsis} Biological Resource Center (ABRC) as DNA "superpools" of 1000 lines each. A primary screen of this population identified insertions in five different HDZip genes, \textit{ATHB5}, 6, 7, 40 and 51. The insertions in \textit{ATHB5}, 7, 40 and 51 most likely result in null mutations (Azpiroz-Leehan and Feldmann, 1997) as these mutation are either located within intron sequences or, in the case of the \textit{ATHB7} insertion, in the 5´untranslated leader sequence. In contrast, a T-DNA inserted upstream of the transcription start site of the \textit{ATHB6} gene might affect gene expression of \textit{ATHB6}, resulting in a "knock-down" mutation (Krysan et al., 1999) of the \textit{ATHB6} gene.

In the case of \textit{ATHB13}, a putative null mutant line was isolated using the same strategy as described above but from a collection of lines harboring \textit{En-1} transposons (Baumann et al., 1998). This insertion was located in the 5´ splice site of the first intron of \textit{ATHB13} and most likely results in a null-mutation of the \textit{ATHB13} gene.

The mRNA and protein levels of the gene have to be assessed in order to confirm that the insertion has created a null-mutation. This was done for the \textit{ATHB5} insertion line and the result showed that there was no detectable \textit{ATHB5} mRNA or proteins in extracts of this line, confirming that this insertion in the first intron of the \textit{ATHB5} gene results in a null-mutation. To what extent the insertions in \textit{ATHB6}, 7, 40 and 51 disrupt gene transcription and translation awaits further testing.
When analyzing the phenotype of the insertion lines it became evident that none of these lines exhibited any visible phenotypic deviations from wild type, neither during seedling growth nor as adult plants. The insertion lines of *ATHB5* and *I3* were subjected to more thorough phenotypic analyses with respect to ABA and sucrose sensitivity, respectively. In neither case did the loss-of-function mutations confer any phenotypic deviations as compared to the wild type. One explanation to the lack of distinguishable phenotypes in the insertion lines might be that HDZip genes are only functional under specific physiological conditions, i.e. unless the mutant plant is subjected to conditions in which the gene is essential, no phenotypic deviations from wild-type will be visible. Another explanation is gene redundancy, i.e. the loss of gene activity in these insertion lines might be compensated by other HDZip genes. Considering the analysis of DNA-binding preferences of HDZip proteins it is likely that some of these proteins interact with identical binding sites and activate the same target genes *in vivo*.

In summary, the use of a PCR-based systematic approach for mutant isolation in *Arabidopsis* HDZip I genes proved to be successful; from a population of 6000 insertion lines, T-DNA insertions in five different genes were identified. The fact that these lines lack distinguishable phenotypes emphasizes the importance of establishing careful screens for growth conditions in which the activity of the target genes are required. Alternatively, for those genes that act redundantly, genetic crosses between plants that bear mutations in these genes have to be performed.
SUMMARY

HDZip proteins constitute a large family of transcription factors in plants and with the close completion of the *Arabidopsis* genome sequencing, we estimate that the total number of HDZip I and II genes will not exceed 26. HDZip proteins interact specifically with pseudopalindromic binding sites, in which each half-site is occupied by one HDZip monomer. The difference in binding specificity between HDZip proteins resides primarily in their preference for the central position of the binding site, whereas the interaction with each half-site appears to be conserved. The finding that HDZip I proteins are capable of forming both homo- and heterodimers *in vitro* suggests that HDZip proteins might be involved in a complex network of interacting transcription factors in the plant.

The HDZip I protein ATHB5 was shown to be active in the regulation of seed germination and postgerminative growth, possibly as an ABA-dependent regulator of cell expansion. ATHB13 affects leaf morphology as well as gene regulation in response to sucrose, suggesting that ATHB13 takes part in a sucrose signaling pathway. The functional analysis of these two proteins was primarily based on the interpretation of overexpression phenotypes. Establishment of HDZip I gene function based on phenotypic analyses of mutant plants identified by reverse genetics screens from collections of T-DNA insertions and transposable elements has as yet not been successful. The reason for this is not clear; however, it is possible that HDZip genes are redundant or that the conditions in which HDZip genes are required are not known.
ACKNOWLEDGEMENTS

I wish express to express my sincere thanks to all those who have contributed to this thesis:

My supervisor, Professor Peter Engström, for accepting me as PhD-student and teaching me scientific thinking and writing.

Johannes, for interesting discussions about the deeper secrets of HDZip science; invaluable help with computer-problems, and for giving me the opportunity to eat fast food every now and then.

Thanks are due to Eva Söderman for being the perfect room-mate.

Eva Söderman, Eva Sundberg and Johannes Hanson, for critical reviewing parts of this thesis.

The members of the Prof lab, Yan, Mats and Jens, for providing interesting discussions and good friendship during labwork; Mats and Jens for making the Prof lab the most funky lab in the building. Thanks to Peter, again, for letting Prof lab stay funky!

I wish to thank all people that have contributed to this thesis by technical assistance: Marie Lindersson, Marie Englund, Eva Burén, Kajsa Pöntinen, and Agneta Ottoson.

Birgitta, thank you for helping me with all imaginable administrational problems.

Thank you Ingela for nostalgic conversations about good old times! Thank you Anneli, Sandra, Katarina, and the rest of cellskapet for making fysbot a nice place to work in!

Last, but not least, I would like to thank:

Hanna for love and support, but also for invaluable comments on all manuscripts of this thesis.

Svante, for being the sunshine of my life!

My parents and Hannas parents for continuos support and for taking care of Svante during the work with this thesis.
REFERENCES


Bellman, R. and Werr, W. 1992. Zmho1a, the product of a novel maize homeobox gene, interacts with the *Shrunken* 26 bp feedback control element. EMBO J. 11: 3367-74


Mattsson J. 1995. Homeobox genes in plants. Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology, 140 Uppsala, Acta Universitas Upsaliensis.


