MOLECULAR MECHANISM OF ARABIDOPSIS CBF1 MEDIATED PLANT COLD-REGULATED GENE TRANSCRIPTIONAL ACTIVATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the Graduate

School of The Ohio State University

By

Zhibin Wang, M.S.

*****

The Ohio State University

2006

Dissertation Committee:

Professor Eric J. Stockinger, Advisor

Professor Michelle Jones

Professor Sophien Kamoun

Professor David Mackey

Approved by

Horticulture and Crop Science Graduate Program
ABSTRACT

Low temperature, drought, and salinity, which eventually result in dehydration, are the most important environmental stresses limiting the planting area and causing devastating yield losses of many crop species worldwide. Over the past century, traditional breeding gained limited success in enhancing crop stress tolerance. Thus, understanding the molecular mechanisms of plant stress tolerance is required to provide an effective strategy to improve crop stress tolerance.

To elucidate the molecular mechanisms of plant cold acclimation, I focused on the understanding of Arabidopsis CBF1 (C-repeat binding factor 1, CRT) mediated plant cold-regulated (COR) gene trans-activation, using a combination of bioinformatic, biochemical, genetic, and molecular biology methods. The work was initiated 1) to mutationally analyze the CBF1 activation domain (AD), 2) and then to determine the epigenetic regulation on COR gene that is played by the putative associated chromatin-modifying factors of CBF1AD.

Mutagenesis strategies including alanine-scanning mutagenesis, truncation mutation, and leucine substitution were used to determine the functional roles of amino acids within the CBF1 COOH-terminus. I created all these mutants as fusions to the heterologous GAL4 DNA binding domain. Mutants were then transformed into
yeast reporter strain harboring a GAL4 responsive reporter construct, consisting of a β-galactosidase gene driven by a promoter with four GAL4 binding sites, or a GAL4-responsive GUS gene Arabidopsis reporter line. The trans-activity of mutants was then determined by the assay of reporter gene. Truncation mutation and alanine substitution analyses suggested there were multiple regions contributing to trans-activation. Using the computer program Hydrophobic Cluster Analysis (HCA), I identified several hydrophobic clusters (HC) composed of bulky hydrophobic or aromatic amino acids. The disruption of more than two HCs was required to significantly reduce the transactivation activity, suggesting a functional redundancy between HCs. Alignments of Arabidopsis CBFs with homologs from different species reveal that the hydrophobic patterns in Arabidopsis CBF1 are conserved across plant taxa. Further leucine substitution of the conserved alanine residues suggested that the residue size contributed to the CBF1 maximal trans-activity. Alanine substitution of the most conserved WSY motif revealed an inhibitory effect on the trans-activity. Furthermore, my mutagenesis also suggests that the growth retardation resulted from the overexpression of a trans-activator in Arabidopsis is related to the trans-activity of that trans-activator. Thus, the C-terminus of CBF1 contains conserved amino acid stretches with special functional roles, including alanine stretches, hydrophobic cluster, and WSY motif. My extensive analyses provide an example that how a plant transcription factor adopts a highly ordered structure to fine-tune gene expression.

The second part of this dissertation is to determine the role of epigenetic control of transcription in the regulation of COR gene expression. Several methylation mutant lines were subjected to low temperature in a time course treatment. RNA gel blot analyses
revealed that a single T-DNA integration locus with multiple copies of rearranged transgene in these methylation mutants resulted in the reduction of *CBF1*, *COR15a*, and *ABF1* gene expression, suggesting a link between the integrated repeats and transcriptional regulation of genes involved in plant cold acclimation. Plant freezing assays confirmed that the transgene locus affected the plant freezing tolerance.

Next, I determined the putative direct target of CBF1 activation domain, the homologs of yeast Tra1 (transcription domain-associated protein 1), and its role in regulation of *COR* gene expression. Tra1 is the only common and critical component of histone acetyltransferase complexes involved in chromatin-modifying that facilitate transcriptional activation. I show that Arabidopsis encodes two Tra1-like genes with high identity and designate them as *AtTra1-2* and *AtTra1-4*. Protein sequence analyses revealed that AtTra1 proteins contained multiple protein-protein interaction motifs with putative function in transcriptional regulation. Homozygous T-DNA insertion mutants in *AtTra1* had no obvious phenotype under normal growth conditions. However, the double homozygous T-DNA mutants appeared to be lethal, indicating that AtTra1 was essential and that two genes might be functionally redundant. Mutants with heterozygosity in one *AtTra1* gene and homozygosity in another *AtTra1* gene develop normally, but siliques contain aborted embryos, which indicated that lethality occurs sometime during embryogenesis. RT-PCR analyses suggested that T-DNA insertions resulted in truncated transcripts that were lack the conserved PI-3 kinase (PI3K) domain. RNAi mediated knockdown expression of AtTra1 resulted in pleiotropic phenotypes on plant growth and development, including elongated hypocotyls, short statue, smaller and abnormally colored cotyledons. *AtTra1-2* promoter::GUS fusion constructs indicated that AtTra1-2
is expressed in meristem and mitotically active tissues. Thus I conclude that AtTra1
genes are essential to Arabidopsis embryogenesis, growth and development.

This research provides insight into how a plant transcriptional activator regulates
gene expression, and the mechanism of plant cold acclimation. The revealed relation
between the trans-activity and the resulting growth stunted phenotype is helpful to design
new strategies for the application of transcription factors in agricultural biotechnology.
Dedicated to my wife Xiaotang Jiang,

my daughter Yijia, and my parents
ACKNOWLEDGMENTS

I wish to offer special thanks to my advisor Dr. Eric J. Stockinger, for his intellectual support, encouragement, and enthusiasm throughout my degree program, and for the freedom to do any experiment I am interested. His great patience always makes me feel free to ask lots of questions. Without his sincere support, guidance, and help, this work would not have been possible. I am forever grateful to him. I also wish to thank my other committee members, Dr. Michelle Jones, Dr. Iris Meier, Dr. David Mackey, and Dr. David Francis for their assistance and invaluable advice. Their encouragements energize me to sail the “scientific ocean” with new hopes to re-, re-, and re-search the “scientific jewelries”. I need also to offer my very special thanks to Dr Sophien Kamoun for being my dissertation member.

Many thanks to the members of Stockinger lab, Stephanie Roberts, Dr Joyce Pennycook, Annie Knox, and Kip Gardener for their assistance during my studies. I also would like to thank Drs Wencai Yang, Xiaodong Bai, Han Xiao, Miaoying Tian for the stimulating discussions during my scientific research. The Ohio Agricultural Research and Developmental Center (OARDC) Graduate Studies Enhancement Grant funded part of my research. I should acknowledge that the Department of Horticulture and Crop Science and National Science Foundation (NSF) supported me with salaries, tuitions and fees. Two Travel Grants, one from American Society of Plant Biologists and one from
Society for Developmental Biology, supported me to present my results at 2006 annual meetings.

Finally, I want to thank my wife Xiaotang Jiang and my daughter Ashley Yijia Wang for their great patience and understanding of my days, nights, and weekends in the laboratory. I need also thank my parents (-in-law), brothers and sister for their financial supports and spiritual encouragements. Without their supports, this goal would not be possible to attain.
VITA

1992………………………………………. B.S. Laiyang Agricultural College, China

1992-1995………………………………….Analytical chemist, Linyi City Bureau of Town and Township Enterprises, China

1995-1998………………………………….Graduate Research Assistant, Graduate School, Chinese Academy of Agricultural Sciences, Beijing, China

1998…………………………………………M. S. Graduate School, Chinese Academy of Agricultural Sciences, Beijing, China

1998-1999………………………………….Research associate, Biotechnology Research Institute (Center), Chinese Academy of Agricultural Sciences, Beijing, China

1999-2000………………………………….Assistant professor, Biotechnology Research Institute (Center), Chinese Academy of Agricultural Sciences, Beijing, China

2000-present……………………………Graduate Teaching and Research Associate, The Ohio State University

PUBLICATIONS


in transgenic plants. *Chinese Bulletin of Botany* **17**:108-113


**FIELDS OF STUDY**

Major Field: Horticulture and Crop Science

Specialization: Plant Molecular Biology and Biotechnology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>vi</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vii</td>
</tr>
<tr>
<td>Vita</td>
<td>ix</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xiv</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xv</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Regulation of genes involved in plant cold acclimation</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Overview</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Two signal transduction pathways leads to plant COR gene transcriptional activation</td>
<td>3</td>
</tr>
<tr>
<td>1.2.1 Overview</td>
<td>3</td>
</tr>
<tr>
<td>1.2.2 ABA-independent pathway (CBF) pathway</td>
<td>3</td>
</tr>
<tr>
<td>1.2.3 ABA-dependent pathway</td>
<td>6</td>
</tr>
<tr>
<td>1.3 Structure and function of CBF transcription factors, a small subfamily of AP2/ERF (ethylene-responsive element binding factor) transcription factors</td>
<td>7</td>
</tr>
<tr>
<td>1.3.1 Overview</td>
<td>7</td>
</tr>
<tr>
<td>1.3.2 Structure of AP2 domains and binding specificity</td>
<td>9</td>
</tr>
<tr>
<td>1.3.3 Activation domain structure and regulation of gene expression</td>
<td>11</td>
</tr>
<tr>
<td>1.3.4 Activation domain and growth stunting</td>
<td>14</td>
</tr>
<tr>
<td>1.3.5 Repression domain and dominant-negative effect</td>
<td>16</td>
</tr>
<tr>
<td>1.3.6 Transcription factors and applications in biotechnology</td>
<td>16</td>
</tr>
<tr>
<td>1.3.7 Conclusion remarks</td>
<td>18</td>
</tr>
<tr>
<td>1.4 Chromatin-remodeling</td>
<td>19</td>
</tr>
<tr>
<td>1.4.1 Overview</td>
<td>19</td>
</tr>
<tr>
<td>1.4.2 SWI/SNF complex and histone-DNA disruption</td>
<td>20</td>
</tr>
<tr>
<td>1.4.3 SAGA complex and histone acetylation</td>
<td>22</td>
</tr>
</tbody>
</table>

xi
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4.4 SAGA complex and environmental stresses</td>
<td>23</td>
</tr>
<tr>
<td>1.5 Acknowledgments</td>
<td>24</td>
</tr>
<tr>
<td>1.6 References</td>
<td>25</td>
</tr>
<tr>
<td>2. Multiple hydrophobic motifs in Arabidopsis CBF1 activation domain</td>
<td>36</td>
</tr>
<tr>
<td>2.1 Abstract</td>
<td>36</td>
</tr>
<tr>
<td>2.2 Introduction</td>
<td>37</td>
</tr>
<tr>
<td>2.3 Materials and methods</td>
<td>41</td>
</tr>
<tr>
<td>2.4 Results</td>
<td>46</td>
</tr>
<tr>
<td>2.5 Discussion</td>
<td>53</td>
</tr>
<tr>
<td>2.6 Acknowledgments</td>
<td>57</td>
</tr>
<tr>
<td>2.7 References</td>
<td>57</td>
</tr>
<tr>
<td>3. Extensive mutagenesis reveals a highly ordered structure</td>
<td>69</td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>69</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>70</td>
</tr>
<tr>
<td>3.3 Materials and methods</td>
<td>73</td>
</tr>
<tr>
<td>3.4 Results</td>
<td>75</td>
</tr>
<tr>
<td>3.5 Discussion</td>
<td>78</td>
</tr>
<tr>
<td>3.6 Acknowledgments</td>
<td>84</td>
</tr>
<tr>
<td>3.7 References</td>
<td>85</td>
</tr>
<tr>
<td>4. Crosstalk between pathways involved in low temperature signal</td>
<td>98</td>
</tr>
<tr>
<td>4.1 Abstract</td>
<td>98</td>
</tr>
<tr>
<td>4.2 Introduction</td>
<td>99</td>
</tr>
<tr>
<td>4.3 Materials and methods</td>
<td>102</td>
</tr>
<tr>
<td>4.4 Results</td>
<td>108</td>
</tr>
<tr>
<td>4.5 Discussion</td>
<td>110</td>
</tr>
<tr>
<td>4.6 References</td>
<td>112</td>
</tr>
<tr>
<td>5. Methylation mutants harboring multiple copies of a transgene within</td>
<td>122</td>
</tr>
<tr>
<td>5.1 Abstract</td>
<td>122</td>
</tr>
<tr>
<td>5.2 Introduction</td>
<td>123</td>
</tr>
<tr>
<td>5.3 Materials and methods</td>
<td>128</td>
</tr>
<tr>
<td>5.4 Results</td>
<td>131</td>
</tr>
<tr>
<td>5.5 Discussion</td>
<td>134</td>
</tr>
<tr>
<td>5.6 Acknowledgments</td>
<td>139</td>
</tr>
</tbody>
</table>

xii
5.7 References.............................................................139

6. Arabidopsis homologs of transcription domain-associated protein 1 (Tra1) are essential for embryogenesis, growth, and development................................................149

6.1 Abstract ...............................................................149
6.2 Introduction ..........................................................150
6.3 Materials and methods ............................................154
6.4 Results ....................................................................161
6.5 Discussion ..................................................................172
6.6 Acknowledgments ....................................................181
6.7 References..............................................................181

Bibliography ......................................................................212
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1  Summary of hydrophobic clusters with transcriptional activities</td>
<td>89</td>
</tr>
<tr>
<td>5.1  Update of characteristics of plant TGS in trans</td>
<td>144</td>
</tr>
<tr>
<td>6.1  Oligonucleotides used as PCR primers in the text</td>
<td>188</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Alignment of the C-terminal 98 amino acid sequence (116-213) of Arabidopsis CBF1 with that of CBF2 and CBF3</td>
<td>62</td>
</tr>
<tr>
<td>2.2 β-galactosidase activities of truncation and alanine substitution mutants</td>
<td>63</td>
</tr>
<tr>
<td>2.3 Hydrophobic Cluster Analysis (HCA) of CBF1 carboxy terminal 101 amino acid residues revealed six hydrophobic cluster (HC) motifs</td>
<td>64</td>
</tr>
<tr>
<td>2.4 β-galactosidase activities of hydrophobic cluster mutants</td>
<td>65</td>
</tr>
<tr>
<td>2.5 RNA gel blot analysis indicates that hydrophobic clusters critical for CBF1 activation domain transactivation in yeast are also critical in Arabidopsis</td>
<td>66</td>
</tr>
<tr>
<td>2.6 ClustalX alignment of the C-terminal end of Arabidopsis CBFs and homologous CBF proteins from crop species</td>
<td>67</td>
</tr>
<tr>
<td>2.7 Non-activating HC2/3/4 mutants exhibit a normal growth phenotype</td>
<td>68</td>
</tr>
<tr>
<td>3.1 The \textit{uidA} expression levels induced by WNH196 and WT</td>
<td>90</td>
</tr>
<tr>
<td>3.2 β-galactosidase levels of individual GMP alanine substitution mutants</td>
<td>91</td>
</tr>
<tr>
<td>3.3 Alignment of the COOH-terminal regions of dicot CBF proteins</td>
<td>92</td>
</tr>
<tr>
<td>3.4 β-galactosidase activity levels, and \textit{uidA} expression levels of WSY mutants</td>
<td>93</td>
</tr>
<tr>
<td>3.5 The \textit{uidA} expression levels of leucine scan mutants in Arabidopsis</td>
<td>94</td>
</tr>
<tr>
<td>3.6 Relative trans-activation activity of the VP16 and CBF1 activation domains</td>
<td>95</td>
</tr>
<tr>
<td>3.7 Predicted phosphorylation sites within CBF1</td>
<td>96</td>
</tr>
<tr>
<td>3.8 “MaWeiBa Gongnong” model for Arabidopsis trans-activator CBF1, adopting a highly ordered structure to fine-tune its downstream gene expression</td>
<td>97</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>4.1</td>
<td>β-galactosidase levels of transcriptional activators in reporter strains harboring different cis-acting elements.</td>
</tr>
<tr>
<td>4.2</td>
<td>Sequences upstream of TATA box in COR15a promoter region that harbors three CRT/DREs and three ABREs were used to construct yeast reporter strain ZW81.</td>
</tr>
<tr>
<td>4.3</td>
<td>Construction of four ABFs expression vectors.</td>
</tr>
<tr>
<td>4.4</td>
<td>β-galactosidase level of ABF and CBF transcriptional activators in yeast reporter strain ZW81 harboring COR15a promoter.</td>
</tr>
<tr>
<td>4.5</td>
<td>Alignment of Arabidopsis ABF1 and rice TRAB1.</td>
</tr>
<tr>
<td>5.1</td>
<td>CBF and COR gene expression analyses in methylation mutants.</td>
</tr>
<tr>
<td>5.2</td>
<td>Freezing assay of mutant lines.</td>
</tr>
<tr>
<td>5.3</td>
<td>A model for multiple copies of HPT gene in one locus epigenetically regulating genes involved in plant cold acclimation.</td>
</tr>
<tr>
<td>6.1</td>
<td>Arabidopsis contains two AtTra1 genes.</td>
</tr>
<tr>
<td>6.2</td>
<td>Schematic illustrations of the AtTra1-2 (A) and AtTra1-4 (B) motifs and domains.</td>
</tr>
<tr>
<td>6.3</td>
<td>LxxLL motifs present in AtTra1-2 and AtTra1-4.</td>
</tr>
<tr>
<td>6.4</td>
<td>Alignment of FATC domains.</td>
</tr>
<tr>
<td>6.5</td>
<td>Molecular characterizations of AtTra1 T-DNA insertion mutants.</td>
</tr>
<tr>
<td>6.6</td>
<td>Expression analyses of CBF1, ABF1 and COR genes in three T-DNA insertion mutants.</td>
</tr>
<tr>
<td>6.7</td>
<td>RT-PCR expression analyses of AtTra1-2 and AtTra1-4 in Arabidopsis tissues.</td>
</tr>
<tr>
<td>6.8</td>
<td>Genotypic analyses of AtTra1-2D x AtTra1-4K F₂ population.</td>
</tr>
<tr>
<td>6.9</td>
<td>Phenotypes of T-DNA insertion mutant and wild type plant.</td>
</tr>
<tr>
<td>6.10</td>
<td>Segregation analysis of embryo lethality associated with AtTra1 mutations.</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>6.11 Disruption of both AtTra1 genes results in embryo lethality</td>
<td>205</td>
</tr>
<tr>
<td>6.12 Alignment of AtTra1-2 and AtTra1-4 genomic regions with poplar cDNA clone PO 02035A02</td>
<td>206</td>
</tr>
<tr>
<td>6.13 β− glucuronidase (GUS) staining of Arabidopsis plants transformed with the AtTra1-2:GUS fusion</td>
<td>208</td>
</tr>
<tr>
<td>6.14 Schematic representation of two RNAi constructs</td>
<td>209</td>
</tr>
<tr>
<td>6.15 The phenotypes of RNAi lines</td>
<td>210</td>
</tr>
<tr>
<td>6.16 Expression analyses of cold inducible genes in RNAi lines</td>
<td>211</td>
</tr>
</tbody>
</table>
CHAPTER 1

REGULATION OF GENES INVOLVED
IN PLANT COLD ACCLIMATION

1.1 Overview

Low temperature is a critical environmental stress that limits the planting area of crop and horticultural species throughout the world. In spring, frosts can kill the style, the female reproductive structure of flowers, causing yield loss of horticultural, crop species in North America. In autumn, early frosts can destroy potential harvests leaving nothing. Every year, low temperature causes crop losses estimated to be $200 million to $1 billion in the US (http://www.mendelbio.com/products/freezing.html).

Drought is another major environmental stress. The financial losses caused by drought are also astonishing with over $10 billion per year globally (http://www.mendelbio.com/products/drought.html). In the US, between 1978 and 1995, the average crop losses caused by drought were more than $1.2 billion each year. Thus, low temperature and drought can cause huge losses each year. However, over the past century, traditional breeding methods had limited success at enhancing tolerance to cold and drought (Thomashow, 1990; Richards, 1996; Fowler, 1999). This is due to several factors: 1) reduced focus on breeding for traits related to stress tolerance rather
than on yield; 2) difficulty in breeding for cold and drought tolerance (Tester and Bacic, 2005). Thus, understanding the molecular mechanism of plant cold tolerance becomes necessary and urgent, in order to provide an alternative strategy to increase plant adaptation to these stresses.

Plants vary in their abilities to tolerate low temperatures. Many plants can significantly increase their cold tolerance after a short exposure to low, non-freezing temperatures, a process known as cold acclimation. For example, rye is killed at –5°C without cold acclimation, but can survive temperatures down to –30°C after cold acclimation (Thomashow, 1999). Plant cold acclimation involves two major pathways, an ABA- (abscisic acid) independent pathway and an ABA-dependent pathway. The ABA-independent pathway leads to the activation of a small family of transcriptional activator proteins identified as the CBFs (C-repeat binding factors). The CBFs appear to be the “master switches” controlling many downstream plant cold-regulated (COR) genes (Stockinger et al., 1997; Sarhan, 1998). The expression of these COR genes is thought to increase plant cold tolerance by protecting plant cells against dehydration, which occurs as a consequence of freezing or drought (Thomashow, 1999). In contrast to the ABA-independent pathway, the ABA-dependent pathway involves the plant hormone abscisic acid (ABA). Exposure of Arabidopsis to low temperatures results in elevated ABA levels, and the expression of ABA-responsive element binding factors (ABFs).

Many questions regarding plant cold acclimation remain unanswered. For example, how do Arabidopsis plants sense temperature decrease and transduce this signal to transactivate CBF genes and ABF genes? How do CBF and ABF transactivate COR
genes? What is the structure of transcription factors like CBF1 and how does that structure permit them to fulfill their functional roles during the adaptation to stresses? What factors interact with these transcription factors to transactivate \textit{COR} genes? What is the relationship between CBF and ABF in transactivation of \textit{COR} genes? In this review, I summarize what is known in the research areas by addressing each of these questions.

1.2 Two Signal Transduction Pathways Leads to Plant \textit{COR} Gene Transcriptional Activation

1.2.1 Overview

It is now clear that plant cold acclimation involves at least two pathways, an ABA-independent pathway and an ABA-dependent pathway. These two pathways involve different cascades of transcription factors. Presently, it is not clear whether these two pathways share the same mechanism to sense the temperature decrease and to transmit the perceived signal to transduce the final activation of plant \textit{COR} genes.

1.2.2 ABA-independent Pathway (CBF Pathway)

Characterization of cold acclimation leads to the identification of \textit{CBF} genes. In 1970, Weiser proposed that changes in gene expression might occur during cold acclimation (Weiser, 1970), and that these gene expression changes might be a key factor in cold acclimation. In 1985, Guy et al. first showed that changes in gene expression occurred during cold acclimation (Guy et al., 1985). Since those first demonstrations by
Guy, many investigators have identified numerous cold regulated \((COR)\) genes (Hajela et al., 1990; Kurkela and Franck, 1990). In some instances, overexpression of these \(COR\) genes enhances plant cold tolerance (Artus et al., 1996; Uemura et al., 1996), supporting the important role played by \(COR\) genes during plant cold tolerance. Functional characterization of the promoter region of \(COR15a\) identified the \(cis\)-acting element CCGAC (C-repeat/Dehydration Responsive Element, CRT/DRE) (Baker et al., 1994; Yamaguchi-Shinozaki and Shinozaki, 1994). Using the yeast one-hybrid system, Stockinger et al. (1997) identified a transcription factor that binds to the CRT, and named this as CRT binding factor 1 (CBF1). CBF1 is a 213 amino acid protein that contains an AP2 domain in its central region and an acidic region in its C-terminus (Stockinger et al., 1997). Overexpression of CBF1 in Arabidopsis induced \(COR\) gene expression in the absence of a low temperature stimulus and resulted in significantly increased plant cold tolerance (Jaglo-Ottosen et al., 1998).

CBF1 belongs to a small subfamily of AP2 transcription factors with six members. The Arabidopsis genome encodes about 146 AP2 domain containing factors according to the Database of Arabidopsis Transcription Factors (DATF) (www.daft.cbi.pku.cn) (Guo et al., 2005). Among them, six proteins have two “signature sequences” PKK/RPAGRxKFxETRHP and DSAWR that flank the AP2 domain (Jaglo et al., 2001). CBF1, CBF2, and CBF3 are involved in plant cold acclimation (Stockinger et al., 1997; Gilmour et al., 1998; Gilmour et al., 2000); CBF4 is involved in drought adaptation (Haake et al., 2002); whereas the functional role of CBF5 and CBF6 (At1g12610 and At1g63030) has not been determined.
Although the components of the CBF regulon are conserved in plant species, CBF regulons differ in activation in response to environmental conditions. Our understanding of the mechanism of cold acclimation advanced quickly through the research in Arabidopsis. Arabidopsis CBF mediates the expression of an entire pathway of genes identified as the CBF regulon. Components of the CBF regulon are also found in other crop species (Jaglo et al., 2001; Zhang et al., 2004). CBF proteins with the signature peptide PKK/RPAGRxKFxETRHP and DSAWR are well conserved in dicots like Arabidopsis, canola, cotton, grape, lettuce, Medicago truncatula, soybean, sunflower, and tomato, and in monocots like barley, maize, wheat, and rye (Jaglo et al., 2001; Zhang et al., 2004; Skinner et al., 2005; Wang et al., 2005). Additionally, the overexpression of Arabidopsis CBF genes in *B. napus* results in the activation of genes harboring the CRT/DRE in their promoters (Jaglo et al., 2001). On the other hand, although different crop species contain conserved CBF regulons, homologous CBFs from different crop species differ in activation in response to environmental stresses. For example, tomato is unable to cold acclimate. Only one of its three homologous genes, *LeCBF1*, *LeCBF2*, and *LeCBF3*, was activated in response to low temperature (Zhang et al., 2004). Furthermore, overexpression of Arabidopsis or tomato CBFs in tomato only induced the expression of four *COR* genes (Zhang et al., 2004). Thus, the components of CBF regulon are conserved among plant species including those that cannot cold acclimate, but differ in regulation of gene expression in response to environmental stimuli.
1.2.3 ABA-dependent Pathway

It has long been recognized that the plant hormone abscisic acid (ABA) plays a major role in adaptation to a wide range of abiotic environmental stresses, including drought, cold, and high salinity (Leung and Giraudat, 1998; Shinozaki and Yamaguchi-Shinozaki, 2000). Environmental stresses induce increases in endogenous ABA levels and activate many genes. Exogenous application of ABA also activates many of these same stress-related genes (Thomashow, 1999). Mutant plants that have either lost the ability to sense ABA or that are impaired in ABA biosynthesis are more sensitive to environmental stresses than wild type plants (Heino, 1990; Mantyla et al., 1995). Thus, ABA plays a critical role in triggering the plant response to these environmental stresses (Leung and Giraudat, 1998).

Functional characterization of the promoters of numerous ABA regulated genes identified a conserved cis-acting element, now referred to as the ABRE (ABA-responsive element). The ABRE is somewhat variable, most often PyACGTGGC, but in general all ABREs have the core sequence ACGT (Leung and Giraudat, 1998). Using the yeast one-hybrid screen, four ABRE-binding factors (ABFs) were isolated (Choi et al., 2000; Uno et al., 2000). ABF1 was induced by low temperature but not drought and salt conditions, whereas ABF4 was induced at a lower level by low temperature compared to its induction by salt and drought treatments. ABF2 and ABF3 are not cold inducible but are induced by salinity (Choi et al., 2000).
1.3 Structure and Function of CBF Transcription Factors, a Small Subfamily of AP2/ERF (Ethylene-responsive Element Binding Factor) Transcription Factors.

1.3.1 Overview

As discussed above, numerous transcription factors are involved in the regulation of gene expression in response to environmental stresses. Generally, transcription factors (TFs) are multidomain proteins (modular) containing a DNA-binding domain (DBD) and an effector domain. The effector domain can be either for activation or repression of target genes (Triezenberg, 1995). Some TFs have domains that function in dimerization and/or ligand binding like nuclear receptors in mammalian systems (Nagy et al., 2006). Through the interaction of a DBD with its specific target DNA sequences, a TF is able to “translate” innate genomic sequence signals into specific biological functions via the activity of its effector domain.

Plants have a striking number of transcription factors (Arabidopsis Genome Initiative, 2000). The most recent annotation of the Arabidopsis genome indicates that 6.5% of 26,207 protein-coding genes encode transcription factors (Haas et al., 2005). The number is much higher than in animals, and may be partially due to the sessile feature of plants in that they have to express more TFs to precisely regulate gene expression in a combinational and concert way in response to various environmental stimuli. The AP2 (APETALA2)/ERF (ethylene-responsive element binding factor) family, also known as ethylene-responsive element binding proteins (EREBP) superfamily, is named for the first identified member, APETALA2 (Riechmann and
Meyerowitz, 1998). The AP2 domain was originally considered to be plant specific (Riechmann and Meyerowitz, 1998). However, several recent investigations showed that the AP2-like domains were present in bacteria, bacteriophage, ciliate, apicomplexans, and transposons as part of endonucleases (Magnani et al., 2004; Wuitschick et al., 2004; Balaji et al., 2005; Wessler, 2005). Thus the AP2 domain have evolved from many independent combinations with mobile endonucleases or recombinases from different sources, which may be the reason for the unusual distribution of the AP2 domain (Balaji et al., 2005; Wessler, 2005).

The AP2/ERF superfamily is the third largest family with 146 members in Arabidopsis (Guo et al., 2005). The AP2/ERF members are involved in a wide range of biological processes, including plant growth and development, and response to various environmental stresses (Riechmann and Meyerowitz, 1998). The AP2 domain-containing proteins are divided into three classes: 1) AP2 subfamily which have two AP2 domains, 2) RAV subfamily which have an AP2 domain and a B3 domain for the binding to a bipartite recognition sequence, and 3) ERF subfamily, which have only one AP2 domain (Riechmann and Meyerowitz, 1998). The ERF subfamily includes the ERF and CBF sub-subfamilies. CBF family members are primarily involved in environmental stresses. ERFs are involved in ethylene signal transduction with a potential role in stress adaptation. Many members have been characterized from structure to function since the first AP2 containing factor was identified. For this section, I centered on the ERF subfamily TFs, with special emphasis on their structure and function, which received less attention in the plant community before. Some themes are derived from research in
yeast, *Drosophila* and human. However, observations that many transcription factors function in both yeast and Arabidopsis, as well as recent studies suggest that most of the concepts discovered in other systems are also applicable to plants (Martinez, 2002).

### 1.3.2 Structure of AP2 Domains and Binding Specificity

To exert a specific regulatory role in response to developmental and environmental cues, the TF’s DBD must bind to a specific target DNA sequences (Ptashne, 2004). The binding specificity of AP2 TFs is encoded in ~60 amino acid residues (Ohme-Takagi and Shinshi, 1995; Allen et al., 1998). AP2 domains adopt a β-sheet structure to bind DNA sequences as described below.

**AP2 domains use β-sheet instead of α-helix to bind the DNA major groove.**

The first elucidated AP2 domain structure was AtERF1. AtERF1 binds to AGCCGCC, known as the GCC-box. The GCC box is conserved in the promoter region of ethylene-inducible pathogenesis-related genes (Ohme-Takagi and Shinshi, 1995). The three dimensional structure of the ~60 amino acid residues of ERF1 comprising the GCC-box binding domain (GBD) was determined either in the absence of, or in the complex with its target DNA by nuclear magnetic resonance (NMR). The GBD structure consists of an α-helix packed parallel to three-stranded anti-parallel β-sheet (Ohme-Takagi and Shinshi, 1995; Allen et al., 1998). AtERF1 contacts the DNA double helix in the major groove via its β-sheet instead of its α-helix. This mode of DNA recognition is unlike other DNA-binding proteins interacting through the α-helix. Arg and Trp resides in the three β-sheets bind to bases within the GCC box and also to the sugar phosphate
backbones (Allen et al., 1998; Hao et al., 2002). Mutational analyses of the binding site
AGCCGCC revealed that the second G, fifth G, and seventh C are most critical for ERF
binding (Hao et al., 1998).

What determines the binding specificity of AP2 domain containing
transcription factors? The CRT/DRE, CCGAC is highly similar to the GCC box
GCCGCC. Mutational analyses confirmed that the 5-bp core sequence CCGAC was the
minimal sequence for CBF1 AP2 domain binding (Hao et al., 2002). Full length CBF1
and AP2 domain fragment (CBF1\textsubscript{37-142}) bind to both the GCC box and the CRT \textit{in vitro}
and in transient assays (Hao et al., 2002). The binding dissociation constant (\(K_d\)) (5.6 ±
0.6 nM) of CBF1 AP2 domain to GCC box is close to the \(K_d\) (2.10 ± 2.0 nM) of the
binding to CRT (Hao et al., 2002). In contrast, the tobacco ERF2 (NtERF2) only binds to
the 6-bp GCCGCC motif but not to CRT. Binding specificity is significantly affected by
NtERF2 residue A14, which is conserved among all homologous ERFs (Hao et al., 1998;
Hao et al., 2002). The corresponding residue in CBFs is a valine. Mutation of NtERF2
A14 to V does not affect the binding to the GCC box but results in a \(K_d\) to CRT/DRE
comparable to that of the CBF1’s \(K_d\) to CRT/DRE (Hao et al., 2002). Thus, the binding
specificity of different AP2 domain containing proteins results from the subtle
differences in the residues of the \(\beta\) sheets. These differences may provide a necessary
mechanism to precisely regulate different signal transduction pathways in response to
various environmental stimuli.
1.3.3 Activation Domain Structure and Regulation of Gene Expression

Activation domain adopts \( \alpha \)-helix after induction. In contrast to most DNA binding domains, transcriptional activation domains (AD) do not have a stable solution structure. ADs appear to only adopt a specific structure after interaction with a target factor (McEwan et al., 1996; Shen et al., 1996; Uesugi et al., 1997; Grossmann et al., 2001; Kumar et al., 2004; Watt et al., 2005). For example, herpes simplex virus transcriptional activator protein VP16 C-terminus (the activation region) has no secondary structure (Uesugi et al., 1997; Grossmann et al., 2001). However, after interacting with target hTAF\(_{II}31\), VP16 AD adopts an \( \alpha \)-helix conformation (Uesugi et al., 1997; Grossmann et al., 2001). This inducible structure also occurs for c-Myc, an important bHLH transcription factor involved in tumorigenesis suppression, and human androgen receptor (AR) (McEwan et al., 1996; Kumar et al., 2004). Adding TATA box binding protein (TBP) to c-Myc induces unstructured c-Myc to adopt an \( \alpha \)-helical conformation (McEwan et al., 1996). Upon the interaction of human antigen receptor AD with the RAP74 subunit of TFIIF complex, an \( \alpha \)-helix structure is induced (Kumar et al., 2004).

Multiple hydrophobic amino acid residues are crucial. Most of our understanding of the mechanisms of transcriptional activation is from mutational analyses (Triezenberg, 1995). The ADs are generally classified as acidic, glutamine-rich, or proline-rich, according to the preponderance of amino acid residues present in the protein (Triezenberg, 1995). Extensive mutational analyses of several acidic ADs, including the human viral protein VP16 (Regier et al., 1993; Sullivan et al., 1998), yeast
Gal4 (Ma and Ptashne, 1987), yeast GCN4 (Drysdale et al., 1995; Jackson et al., 1996), and Arabidopsis CBF1 (Wang et al., 2005), provide insight into their functions. From these studies, we know (1) that hydrophobic amino acid residues are critical for trans-activation; (2) that multiple hydrophobic amino acid residues cluster together as independent activation motifs (Regier et al., 1993; Sullivan et al., 1998; Doring et al., 2000; Stebbins and Triezenberg, 2004; Wang et al., 2005); (3) that there are no obvious patterns among different ADs in the linear amino acid sequences; (4) and that acidity of the acidic AD is important for attracting the target via long-range electrostatic forces (Ma and Ptashne, 1987; Benuck et al., 1999; Ma, 2004). Based on the NMR chemical shift change, a recent proposed model supported these observations that unstructured VP16\textsubscript{AD} and their partners were attracted by electrostatic interactions, which were presumably resulted from acidic amino acids, and then VP16\textsubscript{AD} adopted $\alpha$–helix around the important hydrophobic stretches upon interaction with its target coactivator PC4 (positive cofactor 4). This model also explains how VP16\textsubscript{AD} can provide multiple, distinct binding surfaces for multiple interaction targets (Jonker et al., 2005). Alignments of CBFs from multiple species shows that the acidic residues are highly conserved, consistent with an important functional role (Wang et al., 2005).

**Multiple hydrophobic segments in acidic AD acting redundantly and independently in transcriptional activators has its advantages.** Different activation motifs are thought to interact with and recruit different protein factors to the promoter region, which ensures activation of downstream genes during stress conditions (Drysdale et al., 1998; Stebbins and Triezenberg, 2004), and also provides transcriptional activation
synergy (Langdon and Hochschild, 1999). These AD targets may reside in the same or a different coactivator complex. Yeast Hap4 contains two ADs, an N-terminal and a C-terminal AD. Hap4 N-terminal AD is dependent on the Gcn5 coactivator, whereas the C-terminal AD is independent of Gcn5 (Stebbins and Triezenberg, 2004), suggesting different cofactors recruited by two ADs. Thus Gcn5 may provide histone acetylation activity to the promoter region for both AD, thereby making chromatin accessible to general transcription factors. The yeast GCN4 AD contains seven hydrophobic clusters, each of which contributes an additive effect on transcriptional activation \textit{in vivo} (Drysdale et al., 1995). The binding of many subunits, which belong to TFIID and SAGA (Spt-Ada-Gcn5-acetyltransferase) complexes, was completely dependent on these hydrophobic motifs (Drysdale et al., 1998). Secondly, multiple activation motifs could provide insurance for transactivation of stress related genes in a harsh environment (Wang et al., 2005). The consequence of damaging part of the helix could be compensated by other activation motif to interact with partners (Jonker et al., 2005). Thirdly, multiple motifs could add up to be a much more potent transcriptional activator (Drysdale et al., 1998; Wang et al., 2005). VP16 is a potent activator and has two activation subregions (Sullivan et al., 1998). However, CBF1 contains four hydrophobic motifs which together are more powerful than VP16 when assayed as a GAL4\textsubscript{DBD} fusion in Arabidopsis (Wang et al., 2005). The combination of alanine substitutions in hydrophobic clusters of GCN4 resulted in a greater reduction in binding with subunits of complexes, which correlated with their additive effect of trans-activation of GCN4 \textit{in vivo} (Drysdale et al., 1998). This is consistent with earlier observations that trans-activity...
is approximately proportional to the length of Gal4 fragments (Wu et al., 1996; Ptashne and Gann, 1997).

ADs are able to interact with multiple protein factors but may have preferences. ADs are regulatory modules playing their roles by recruiting general transcription factors and stimulating gene transcription. Functional analyses indicated that short peptides of 11 or even fewer amino acids are capable of transactivation (Carpenter et al., 2005). This discovery suggested that activation sequences might not have any enzymatic activity (Ma and Ptashne, 1987), rather their function might be to simply bind other proteins (Ma, 2004). Ma proposed a hypothesis that may explain the seemingly contradictory results between an activating motif and their many different targets (Ma, 2004). First, acidic activating peptides tend to have the potential to bind many target factors in the transcriptional apparatus, because they form multiple distinct binding surfaces when interacting with different partners (Jonker et al., 2005). Second, different acidic activation peptides tend to bind the same target factors, which makes it possible for transcriptional machinery to integrate various activating activities by all possible combinations when needed in response to different developmental and environmental stimuli (Brown et al., 1998; Ma, 2004; Jonker et al., 2005).

1.3.4 Activation Domain and Growth Stunting

Overexpression of TFs in crop plants has the potential to improve important crop traits like stress tolerance. TF overexpression can activate a battery of genes to be expressed (Jaglo-Ottosen et al., 1998; Kang et al., 2002; Kim et al., 2004; Oh et al.,
2005). However, the overexpression of TF in plants sometimes causes stunted growth (Jaglo-Ottosen et al., 1998; Gilmour et al., 2000; Hsieh et al., 2002). This growth stunting would be detrimental to crop yield. Thus, understanding the underlying mechanism causing the stunted growth could provide strategies to overcome this problem while attaining the trait improvements.

Constitutive overexpression of CBF$1_{\text{AD}}$ fused to the heterologous DNA binding domain of GAL4 also resulted in growth stunting without the activation of stress-related genes (Wang et al., 2005). This result argues against the notion that inappropriate activation of many stress-related genes by overexpressed TFs causes growth stunting. The severity of the dwarf phenotype is also linked to the CBF$1_{\text{AD}}$ potency, as partial diminution of activation lessened the severity of the dwarf phenotype (Wang et al., 2005). Similarly overexpression of the less potent activator VP16 resulted in plants with a less severe growth stunting phenotype (Wang et al., 2005). Higher-level expression of TF resulted in more molecules of TF, presumably having similar effect with strong AD. Consistent with this notion, ABF4 expression level determined the degree of retardation in different transgenic Arabidopsis lines (Kang et al., 2002). Overexpressed activators may recruit most of the coactivators in the cell, thus interfering with the transcriptional regulation of housekeeping genes. This hypothesis is called quenching (Ptashne, 1988). To avoid the stunting problem, one feasible alternative strategy is to employ stress-inducible promoters to drive CBF genes, which enhances stress tolerance without growth retardation (Kasuga et al., 1999; Lee, 2003).
1.3.5 Repression Domain and Dominant-negative Effect

Compared to ADs, repression domains (RDs) have been less studied. One reason is that there is no good system to indicate the repression. Within the AP2/ERF superfamily, most members are transcriptional activators. However, some ERF proteins were found to repress target gene expression (Fujimoto et al., 2000; Ohta et al., 2001; Yang et al., 2005). ERF repressors including ERF3, ERF4, and ERFs 7-ERF12 are active repressors, because they mediate repression via their RDs rather than binding to cis-acting DNA regulatory motif and preventing the binding of TFs (Hanna-Rose and Hansen, 1996; Yang et al., 2005). ERF7 is a repressor involved in abscisic acid and drought responses and interacts with AtSin3, a global repressor protein. AtSin3 also interacts with the histone deacetylase HDA19 (Song et al., 2005). Thus AtSin3 may function as a co-repressor to recruit deacetylation activity to the promoter region, preventing the formation of the Pre-initiation complex (PIC) for transcription. Mutational analyses of ERF repressors identified the repression motif $^{L/F}D_{f}L_{f}/(X)P$, named ERF-associated amphiphilic repression (EAR) motif. The EAR motif in ERF7 is thought to be responsible for the interaction with AtSin3 (Ohta et al., 2001). Fusion of the EAR motif to several different transcriptional activators converts these proteins into repressors (Fujimoto et al., 2000).

1.3.6 Transcription Factors and their Applications in Biotechnology

The modular nature of TFs makes it possible to use dominant-negative effect, namely a mutant protein interfering with the function of the wild type protein, to regulate
gene expression both at the lab bench and in the field. Furthermore, plant TFs often belong to a family with many members. For example, the CBF family has six members with apparent functional redundancy among CBF1, CBF2, and CBF3 (Gilmour et al., 2004). This presents a challenge to conduct functional analysis using a reverse genetics strategy to identify loss-of-function alleles. The success of using just the DBD as a dominant-negative strategy however may depend on the chromatin context where the TF’s cognate binding site resides, because it is essential that the DBD binds to the regulatory motif. Chromatin decondensation activity and transcriptional activity of the acidic activating motif can be separated (Carpenter et al., 2005). Thus, fusing a DBD with a peptide only having chromatin decondensation activity may increase the chance of success by helping the DBD binding. Repression domains such as the Drosophila engrailed repressor domain, is frequently fused to gene specific DBDs in order to do functional analyses in animal systems. Markel et al. fused engrailed to several plant TFs and observed similar phenotypes to the corresponding loss-of-function mutants (Markel et al., 2002). When the VP16AD is fused to the EAR motif, the fusion protein loses all trans-activity (Ohta et al., 2001). Thus, the EAR motif can be used to investigate biological functions of plant TFs.

Genome-wide elucidation of gene function also requires approaches to manipulate gene expression at will. Gene-specific control of gene expression (activation or repression) is necessary to define gene functions with redundancy. Using a dominant-negative effect, the polydactyl zinc finger TF provides an alternative and successful approach to manipulate a given gene expression. The polydactyl zinc finger binds to an
18bp DNA sequence, long enough to be unique within a organism’s genome (Beerli and Barbas, 2002). Current zinc finger TFs show its promising future in regulating a given gene expression (Choo and Isalan, 2000; Beerli and Barbas, 2002; Segal, 2002). Designed zinc finger TF that is targeted to an 18 bp DNA sequence fused with VP16AD which trans-activated reporter gene in transient assay or in stably transformed tobacco plants. With a phloem-specific promoter, an artificial TF activated a reporter gene in vascular tissues (Ordiz et al., 2002). Thus custom designed artificial TFs may be a powerful tool both in basic biology and applied biotechnology.

1.3.7 Conclusion Remarks

TFs are modular structured protein factors that bind to specific gene sequences to exert their regulatory roles. Subtle differences in the AP2 DBD determine the binding specificity. Acidic ADs tend to form α-helixes upon interaction with their target protein factors. This is greatly dependent on the hydrophobic regions. ADs and RDs play regulatory roles by recruiting chromatin-remodeling complexes to the promoter region. ADs use these complexes as adaptor to form the PIC, whereas RDs often prevent PIC formation. ADs and RDs interact with multiple protein subunits to form large macromolecular complexes that stabilize the transcriptional regulatory apparatus.

The mechanism responsible for the regulation of eukaryotic gene expression occurs predominantly at the transcriptional level. Thus the sessile nature of plants may require many more transcription factors than animals to adapt to changing environments. This large number of transcription factors is also a result of redundant family members.
Redundancy complicates genetic attempts to reveal functional roles in a specific pathway. Dominant-negative strategies may provide a feasible approach to determine a TF’s role. Furthermore, plant TFs need to recruit multiple components which belong to different chromatin-remodeling complexes for transcriptional activity, thus it is not surprising to find many new and unique components in plants affecting development and responses to environmental stimuli. Considering the multiple protein interactions between cofactors and AD, effects caused by one chromatin-modifying factor could be minor compared to the effect caused by mutation of a TF in a given pathway.

1.4 Chromatin Remodeling

1.4.1 Overview

In vivo, DNA wraps around histone octamers, two copies of four core histone proteins (H2A, H2B, H3, and H4), to package as nucleosomes. These core histone proteins contain short but conserved N-terminal tails that emanate from the compact central histone folds (Eberharter and Becker, 2002). This packed DNA is generally in a state that prevents gene expression. This repressive state raises two questions. First, how do gene-specific transcriptional activators bind to their target sequences? Second, how do activators overcome the repressive effect of chromatin? Much more is known in regards to the second question. Early data indicated that low temperatures reduced the 5-methylecytosine levels in Arabidopsis DNA, and this reduction coincided with increased transcriptional activity (Burn et al., 1993). This may imply that low temperatures induce chromatin remodeling and make the DNA more accessible to TFs.
Two main chromatin-remodeling complexes recruited by gene-specific activators play a role to alter chromatin structure to facilitate gene expression (Pollard and Peterson, 1998; Fry and Peterson, 2001; Lusser and Kadonaga, 2003). These complexes are the SWI/SNF complex, composed of ATPase containing factors, and the SAGA (Spt-Ada-Gcn5-acetyltransferase) complex with histone acetyltransferase (HAT) activity. The SWI/SNF chromatin-remodeling complex contains 11 known subunits, which locally alters or disrupts the histone-DNA interaction using the energy released from ATP hydrolysis (Vignali et al., 2000; Muchardt and Yaniv, 2001). On the other hand, the conserved N-terminal tails of core histone proteins (H2A, H2B, H3, and H4) are often affected by post-translational modification such as methylation, phosphorylation, and acetylation. The SAGA complex acetylates histone lysine residues to alter the chromatin structure (Eberharter and Becker, 2002). Histone acetylation and deacetylation have long been known to be correlated with gene activation, and repression, respectively (Eberharter and Becker, 2002; He et al., 2003; Pipal et al., 2003). The SWI/SNF and SAGA complexes work together rather than independently to make chromatin accessible to transcriptional apparatus (Pollard and Peterson, 1998; Fry and Peterson, 2001; Geng and Laurent, 2004; Govind et al., 2005). Both complexes are required for the expression of yeast genes induced by stress (Krebs et al., 2000; Huisinga and Pugh, 2004).

1.4.2 SWI/SNF Complex and Histone-DNA Disruption

Gene-specific transcriptional activators recruit multiprotein SWI/SNF complex to alter chromatin to regulate gene expression. The ADs of transcriptional
Activators interact with multiple subunits (Snf5, Swi1 and Swi2/Snf2) of the SWI/SNF complex (Fry and Peterson, 2001; Neely et al., 2002; Lemieux and Gaudreau, 2004). Yeast SWI/SNF is a ~2 Mda protein complex composed of 11 subunits. Five subunits, the SWI1, SWI2/SNF2, SWI3, SNF5, and SNF6 proteins, were isolated using a genetic strategy, while six additional proteins were isolated using a biochemical purification strategy (Pazin and Kadonaga, 1997). Homologous complexes exist in other eukaryotic organisms, including flies and mammals. SNF2 (also known as SWI2) is the catalytic subunit of the SWI/SNF complex that harbors the DNA-dependent ATPase. The conserved V motif within the Swi2/Snf2 helicase/ATPase couples ATP hydrolysis to the chromatin remodeling (Smith and Peterson, 2005). Mutations in any of these swi/snf genes resulted in pleiotropic phenotype but not lethal, which suggested their non-essential role. Whole genome expression analyses revealed that the SWI/SNF complex affects about 3 to 6% of all the yeast genes (Holstege et al., 1998; Sudarsanam et al., 2000). Affected genes are dispersed over the whole genome, suggesting that the complex regulates individual genes rather than large blocks of linked genes (Sudarsanam et al., 2000). The in vivo recruitment of SWI/SNF by the yeast transcriptional activator Gcn4 was dependent on SWI/SNF integrity, possibly through multiple interactions with several subunits (Yoon et al., 2004). Furthermore, Gcn4p recruits SAGA and SWI/SNF coactivator and other mediators interdependently (Govind et al., 2005; Qiu et al., 2005).

The Arabidopsis genome contains about 40 SWI/SNF-like proteins. Of these 40 proteins, only a few have been functionally analyzed. Arabidopsis BSH (bushy phenotype), a homolog of the yeast SNF5 protein, partially complemented the yeast snf5
mutation (Brzeski et al., 1999; Reyes et al., 2002). Arabidopsis also contains four homologs of yeast SWI3, AtSWI3A-C (Sarnowski et al., 2002). AtSWI3A and AtSWI3B form homodimers and heterodimers, and interacted with BSH (SNF5), AtSWI3C, and FCA. FCA is an RNA-binding protein promoting flowering (Sarnowski et al., 2005). AtBRM, a homolog of yeast SNF2, contains an ATPase domain that is necessary for vegetative and reproductive development and interacts with AtSWI3C. AtBRM is a subunit of a 1~2 MDa protein complex (Farrona et al., 2004). Another Arabidopsis snf homolog PIE1 (PHOTOPERIOD-INDEPENDENT EARLY FLOWERING1), acts as a coactivator for FLOWERING LOCUS C (FLC) activation (Noh and Amasino, 2003). Thus Arabidopsis also contains SWI/SNF-like complexes that globally regulate gene expression (Brzeski et al., 1999; Reyes et al., 2002; Sarnowski et al., 2005).

1.4.3 SAGA Complex and Histone Acetylation

The yeast SAGA is a 1.8 MDa multiprotein complex composed of approximately 14 polypeptides, and is required for transcription of stress induced genes (Basehoar et al., 2004; Huisinga and Pugh, 2004; Wu et al., 2004). The recent elucidated three dimensional structure of SAGA complex reveals five modular domains. Four domains (domain I to IV) are almost linearly arranged, with domain I at the bottom and domain IV at the top of complex, whereas domain V is at the tip of the complex and attaches to domain IV (Wu et al., 2004). Three protein subunits, Tra1 (transcription domain associated protein 1), Gcn5 and Spt3, play distinct regulatory roles and locate at domain
I, III, and V, respectively (Wu et al., 2004). The Spt3 subunit mediates TBP-TATA interaction (Yu et al., 2003). The Gcn5 protein is the enzymatic subunit that modifies histones through acetylation. The subunit of Tra1, an ataxia telangiectasia mutated (ATM)-related protein, is the direct target of many transcriptional activator ADs (Bhaumik and Green, 2001; Brown et al., 2001; Naar et al., 2001; Bhaumik et al., 2004; Wu et al., 2004), which by association with Gcn5, brings the HAT activity to promoter regions.

1.4.4 SAGA Complex and Environmental Stresses

The SAGA and TFIID complexes share several common subunits Taf5, Taf6, Taf9, Taf10, and Taf12 (Wu et al., 2004). These two complexes also make overlapping contributions to transcriptional regulation of gene expression. SAGA is responsible for regulating ~10% of the yeast genes, which are largely induced by stresses including heat, oxidation, acidity, carbon or nitrogen starvation (Huisinga and Pugh, 2004). In contrast, the TFIID complex is required for expression of ~90% of yeast genes, primarily the housekeeping genes (Huisinga and Pugh, 2004). Heat stress disrupts the TFIID regulated pathway, but enhances the SAGA regulated pathway of stress-induced genes (Zanton and Pugh, 2004).

From size to structure, the yeast SAGA complex is similar to its human counterpart, TFTC, which indicates a functional and component conservation in both organisms (Wu et al., 2004). This aspect of conservation may be due to the evolutionary pressure which retains the essential core function among organisms (Naar et al., 2001). Based on these
aspects, we reasoned that Arabidopsis transcription factors like CBF1 might depend on similar SAGA complex. Though Arabidopsis may contain novel subunits as not all the components of SAGA complex are conserved in the Arabidopsis genome, the crucial subunits including homologs to yeast Ada and Gcn5, AtAda and AtGcn5, are indeed conserved (Stockinger et al., 2001). To be fully functional in yeast, Arabidopsis transcriptional activator CBF1 required yeast Ada and Gcn5 to trans-activate (Stockinger et al., 2001). CBF1 interacts with AtAda and AtGcn5 both \textit{in vitro} and \textit{in vivo} (Stockinger et al., 2001; Mao et al., 2006). Moreover, AtGcn5 has HAT activity (Stockinger et al., 2001). Knocking out AtGcn5 affects the expression of \(~5\%\) of Arabidopsis genes, similar to the effect of GCN5 mutation in yeast (Vlachonasios et al., 2003). Loss-of-function AtGcn5 mutants delayed the expression of plant \textit{COR} genes (Vlachonasios et al., 2003). We hypothesized that AtTra1 was the direct target for CBF1. Arabidopsis genome encodes two Tra1 like genes (Z. Wang and E. J. Stockinger, unpublished data). Reverse genetic analyses indicate AtTra1 plays roles in regulation of plant development and gene expression (Z, Wang and E. J. Stockinger, unpublished data).

\textbf{1.5 Acknowledgments:}

Work in the authors’ lab was supported by NSF, OSU, OARDC to E.J.S, and an OARDC Graduate Competition Grant to Z. W.
1.6 References


Kumar, R., Betney, R., Li, J., Thompson, E.B., and McEwan, I.J. (2004). Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. Biochemistry 43, 3008-3013.


Segal, D.J. (2002). The use of zinc finger peptides to study the role of specific factor binding sites in the chromatin environment. Methods 26, 76-83.


CHAPTER 2

MULTIPLE HYDROPHOBIC MOTIFS IN ARABIDOPSIS CBF1
ACTIVATION DOMAIN REDUNDANTLY CONTRIBUTE TO
TRANSCRIPTIONAL ACTIVATION

2.1 ABSTRACT

Previous data showed that the C-terminus of Arabidopsis CBF1 functioned as a transcriptional activation domain. To determine amino acids critical for transactivation, we employed two strategies, alanine-scanning mutagenesis and truncation mutation. Using alanine-scanning mutagenesis we systematically changed all non-alanine amino acids in blocks of three to four adjacent into alanines. Truncation mutants were introduced at four positions within the C-terminus of CBF1. We created these mutants as fusions to the heterologous GAL4 DNA binding domain. Mutants were transformed into yeast reporter strain harboring a GAL4 responsive reporter construct, consisting of a β-galactosidase gene driven by a promoter with four GAL4 binding sites. To test the effect on transactivation activity caused by mutation, we measured the β-galactosidase activity. Truncation mutation analyses indicated that transactivation activity required
residues C-terminal to amino acid threonine 158. Alanine substitution mutants were consistent with results from truncation mutants. Moreover, alanine-scanning mutants indicated that there were multiple regions that contributed to transactivation; no single amino acid was critical; acidic amino acids played a minor role in transactivation. Using the computational program Hydrophobic Cluster Analysis (HCA), we identified several hydrophobic clusters composed of bulky hydrophobic or aromatic amino acids. Alanine substitution through these clusters indicated that disruption of more than two were required to significantly reduce the transactivation activity. Overexpression of these mutants in a GAL4-responsive GUS gene Arabidopsis reporter line supported the yeast results. Alignments of Arabidopsis CBFs with homologous sequences from different species reveal that the hydrophobic patterns in Arabidopsis CBF1 are conserved across plant taxa. Different HCs are separated by acidic amino acids or prolines. Overexpression of the GAL4-CBF fusions also resulted in growth retardation. Moreover, non-acting GAL4-CBF mutants did not, which suggests that the growth retardation phenotype may in part be related to the capacity for trans-activation. This discovery provides insight into the mechanism for growth retardation.

2.2 INTRODUCTION

Transcription factors (TFs) are usually modular, frequently containing a DNA binding domain and an activity domain, which either activates (activation domain) or represses (repression domain) gene expression. TFs act by binding to a cognate DNA
sequence within the promoter region of a gene to regulate the expression of that gene by its activity domain. TFs play many critical roles involved in signal transduction pathways.

Plants contain many more transcription factors compared to animals, thought to be a result of their sessile nature in order to cope with, and adapt to changing environments during their lifetimes (Arabidopsis Genome Initiative, 2000). Upon abiotic or biotic stimuli, cascades of transcription factors are expressed in appropriate time and space to regulate gene expression in response to the combined external and internal cues (Thomashow, 1999; Xiong et al., 2002). Although plant transcription factors play important roles in the life cycle, little work has been done to biochemically analyze the structure and function of transcription factors’ activity domain.

Traditionally, eukaryotic transcriptional activators are classified as acidic, glutamine-rich, proline-rich, or serine/threonine-rich based on the abundance of the amino acids in the activation domain (AD) (Triezenberg, 1995). However, these are artificial classifications because the abundant amino acids do not necessarily play the critical role in transactivation. For example, the herpes simplex virion protein VP16 is an acidic transcription activator, but it is the aromatic amino acids and bulky hydrophobic amino acid residues that play important roles for transcriptional activation (Regier et al., 1993; Sullivan et al., 1998).

It has become clear that transcription factors have no stable three-dimensional structure when in solution. However, unstructured activation domain adopts an stable α-helical conformation after interacting with their associated protein factors (Shen et al.,
Thus far the most useful method to address the structure and function of an activation domain is mutational analyses, including random mutagenesis and site-directed mutagenesis (alanine substitution mutagenesis). Random mutagenesis is usually conducted through techniques like chemical mutagenesis or error-prone PCR-mediated mutagenesis (Sullivan et al., 1998). Mutants are then selected based on some sort of phenotypic difference. Random mutagenesis allows the investigation of size, charge and hydrophobicity of substituted amino acid sidechains by changing a specific amino acid into any different amino acid. Site-directed mutagenesis allows systematically testing every amino acid and the contribution of their sidechains (Sullivan et al., 1998). Alanine substitution would eliminate the side chain beyond the β-carbon, whereas it would not change the main-chain conformation nor does it cause extreme electrostatic or steric effect (Lefevre et al., 1997).

Arabidopsis CBF1 (C-repeat binding factor 1) belongs to the acidic family of transcriptional activators and activates plant cold-regulated gene expression (Stockinger et al., 1997). Arabidopsis CBF1 was originally identified through a yeast one hybrid strategy by virtue of its ability to bind the cis-acting element CCGAC, referred to as the C-repeat/Dehydration Response Element (CRT/DRE), and transactivate a reporter gene (Stockinger et al., 1997). Subsequently, two additional members CBF2 and CBF3 were identified and shown to be involved in low temperature tolerance (Gilmour et al., 1998; Gilmour et al., 2000). A fourth member, CBF4 was shown to be important in the response to drought (Haake et al., 2002). Under normal growth conditions, CBFs are not
expressed. CBF expression is induced by low temperature, drought, salinity or agitation. Overexpression of *CBF1* in Arabidopsis results in activation of a set of plant cold-regulated genes (*COR*) harboring the C-repeat without low temperature treatment (Jaglo-Ottosen et al., 1998). Thus *CBF1* has the ability to transactivate gene expression as long as it is present in plant cells. CBF overexpressing lines was significantly enhanced in cold tolerance (Jaglo-Ottosen et al., 1998). However, the ectopic overexpression of *CBF1*, *CBF2*, and *CBF3* also results in growth retardation (Jaglo-Ottosen et al., 1998; Gilmour et al., 2004).

*CBF* homologous sequences are present in other plant species, both cold tolerant and cold sensitive (Jaglo et al., 2001; Zhang et al., 2004). Overexpression of *LeCBF1* or Arabidopsis *CBF3* in cold sensitive tomato did not enhance freezing tolerance of tomato. However, overexpression of *LeCBF1* in Arabidopsis activates *COR* genes like *COR15a* and *COR6.6*, whereas only four genes out of 8700 tomato genes were activated in tomato (Zhang et al., 2004). Thus, although different plant species contain homologous *CBF* genes, their innate abilities to activate plant *COR* genes are varied among plant species either cold sensitive or cold tolerant.

Understanding Arabidopsis *CBF* mediated plant *COR* gene transcriptional activation is not only of fundamental importance but is also of practical importance. Over the past century, traditional breeding has gained little success in increasing stress tolerance partially due to the difficulties in identifying clear traits in stress tolerance and less emphasis on stress tolerance selection (Fowler, 1999). Thus, understanding the molecular mechanism for cold tolerance is critical for agriculture to provide an efficient
strategy to improve crop cold tolerance. To understand the molecular mechanism of CBF mediated plant COR gene activation, we started by addressing the structure and function of CBF1 activation domain through extensive mutagenesis. Here we present our data indicating that Arabidopsis CBF1 activation domain contains multiple motifs with redundant contribution to transactivation capacity. These motifs consist of bulky hydrophobic amino acids or aromatic amino acids. We also provide insight about the mechanism for growth retardation, a phenomenon associated with overexpression of many transcriptional activators in plants that potentially has negative consequences for crop yield.

2.3 MATERIALS AND METHODS

Site Directed Mutagenesis

Alanine scan mutagenesis was conducted using primers that incorporated mutations at the target amino acids (Li and Wilkinson, 1997) and using the QuickChange strategy (Stratagene, La Jolla, CA). All mutations were introduced into a GAL4DBDCBF1AD fusion that was in plasmid pGEM11Z- (Promega, Madison, WI). After confirmation by sequencing, the GAL4DBDCBF1ALANINE MUTANTS were subcloned into yeast expression vector pDB20.1 (Berger et al., 1992), and the vector pCIB710 (Rothstein et al., 1987). Followed by KpnI/Hind III digestion to recover the 35S promoter- GAL4DBDCBFALANINE MUTANTS-Terminator cassette, which was then ligated into the Agrobacterium tumefaciens binary vector pCIT20 (Ma et al., 1992).
**Yeast β− galactosidase Assays**

Yeast reporter strain Y190 was used for all quantitative assays and was obtained from the Arabidopsis Biological Resource Center (The Ohio State University, Columbus, OH). The genotype of Y190 (Harper et al., 1993) is MATa, ura3–52, his3–200, ade2–101, lys2–801, trp1–901, leu2–3, 112, tyr1–501, gal4D, gal80D, cyhR2, LYS2::GAL1UAS–HIS3TATA–HIS3, URA3::GAL1UAS-HIS3TATA–lacZ}. Y190 harbors a chromosomally integrated copy of the native GAL1/GAL10 promoter {consisting of four slightly different Upstream Activation Sequence (UAS) dimer binding sites} fused to the lacZ reporter gene (Giniger et al., 1985). For β-galactosidase assays, five independently transformed yeast colonies were inoculated into 5 ml YPAD. Cultures were grown overnight. The next day 200 µl of the overnight cultures were transferred to a new tube containing 10 ml YPAD and grown until the OD_{600} was 0.8-1. Then 5 ml of the cultured cells were harvested and resuspend in 250 µl of breaking buffer {100mM Tris-HCl (pH 8), 1mM Dithiothreitol (DTT), 20% Glycerol} with the addition of 12.5 µl PMSF stock solution (Phenylmethylsulfonyl fluoride, PMSF 100 mM dissolved in 2-propanol). Glass beads (Sigma G8772, acid-washed, 435-600 mm glass beads) were then added. Yeast cells were broken open by vortexing four times at top speed in 20-second bursts after the adding of glass beads. An additional 250 µl breaking buffer was added after the fourth vortexing. Liquid extracts were then transferred to a
new microfuge tube and spun at 13000 rpm for 1 min. Clarified supernatants were transferred to a new tube.

For β−galactosidase assay, 20 µl of the extract was added to 0.9 ml of Z buffer (Na$_2$HPO$_4$.7H$_2$O 60 mM, NaH$_2$PO$_4$.H$_2$O 40 mM, KCl 10 mM, MgSO$_4$.7H$_2$O 1 mM, pH 7, adding 2-ME to 50 nM just before use). Final volume was brought to 1 ml with breaking buffer and then incubated at 28°C water bath for 5 minutes. To initiate the reaction, 0.2 ml of ONPG stock solution (o-nitropheryl-β−D-galactoside, 4 mg/ml in Z buffer) was added and incubated for about 25 minutes. The reaction was terminated with 0.5 ml Na$_2$CO$_3$ solution (1 M). The optical density at 420 nm was measured and the specific activity of the extract was expressed according to the following formula: OD$_{420}$ x 1.7/ (0.0045 x protein x extract volume x time). Each assay was repeated multiple times using independent transformed colonies. The protein concentration was determined by Bradford assay (Bradford, 1976).

**Plant Transformation and Overexpression of GAL4$_{DBD}$/CBF1$_{AD}$ Fusions in Arabidopsis**

Binary vector mutant constructs were electroporated into *Agrobacterium* strain GV3101 (Koncz, 1986). Arabidopsis (ecotype WS-2) was transformed using the floral dip method, which consisted of immersion of plants into a 5% sucrose/ 0.05% Silwet-L77 solution containing ~5x10$^8$ cfu/ml of *Agrobacterium* (Clough and Bent, 1998).

GAL4$_{DBD}$, GAL4$_{DBD}$/VP16$_{AD}$, GAL$_{DBD}$/CBF1$_{AD}$ WT and HC mutants in pCIT20 vectors were transformed into the Arabidopsis reporter line with UAS$_G$ 17 mer (x4) 35S(-46)-
uid A reporter line. Transgenic seedlings were screened on GB5 (Gamborg B5) plates supplemented with 20 ~ 35 µg/ml hygromycin and 300 µg/ml Timenton.

RNA Isolation and RNA Gel Blot Analyses

Independent T₁ seedlings, selected on media supplemented with 20 ~ 35 µg/ml hygromycin, were randomly separated into two to four pools, each pool containing 10 or more seedlings. Total RNA was extracted from pooled seedlings using Plant RNeasy Mini Kits (Qiagen Inc., Valencia, CA). Total RNA was fractioned on 1% formaldehyde gel and blotted to Hybond N + (Amersham Biosciences Corp., Piscataway, NJ). Hybridizations (Tm-20°C) were in 50% formamide, 5X sodium citrate, 20mM sodium phosphate (pH 6.8), 1X Denhardt’s, 0.1% sodium dodecyl sulfate, 10% dextran sulfate and 100 µg herring sperm DNA. Filters were washed in moderate stringency (Tm-20°C) (0.2X sodium citrate, 0.05% sodium dodecyl sulfate and 0.01% sodium phyrophosphate). DNA fragments of uidA, GAL4 DBD, and eIF4A were radiolabeled with 32P-nucleotides by random priming using the Random-Prime Labeling Kit (Amersham Pharmacia Biotech Inc., Piscataway, NJ). RNA blotting results were visualized by phosphorimage autoradiography using a Storm 860 PhosphorImager (Amersham, Piscataway, NJ).

Hydrophobic Cluster Analysis

Hydrophobic cluster analysis (HCA) is a program developed to visualize three dimensional (3D) amino acid sequences in two dimensional (2D) space (Lemesle-Varloot et al., 1992; Woodcock et al., 1992; Callebaut et al., 1997). HCA was conducted using a web based server at http://smi.snv.justsieu.fr/hca-form.html. The principle for motivation
developing HCA is that hydrophobic amino acids are not randomly distributed within protein sequences rather they tend to occur in clusters (Gaboriaud et al., 1987). HCA wraps the linear amino acid protein sequence in a classical α− helix around a cylinder. To visualize this structure in 2D space, the cylinder is cut parallel to its axis, unrolled and then laid out as 2D surface on paper (Gaboriaud et al., 1987). To allow a complete visualization of each amino acid’ surrounding environments, adjacent amino acids were duplicated because they were separated by cutting. Hydrophobic amino acids were considered to be different HC (grouped by outlines) if they are separated by four or more non-hydrophobic amino acids or proline. Proline is thought to be a breaker of clusters. Mosaic clusters (regular alternations of hydrophobic and non-hydrophobic amino acids) are displayed as connecting lines.

**Sequence Comparison**

AP2 domain containing sequences were identified by BLAST searches. Only AP2 domain containing protein sequences with the two CBF signatures (Jaglo et al., 2001) were used for ClustalX alignment. This list includes published proteins, AtCBF1 (Stockinger et al., 1997), AtCBF2 and AtCBF3 (Gilmour et al., 1998), AtCBF4 (Haake et al., 2002), At1g63030 and At1g12610 (Arabidopsis Genome Initiative, 2000), BnCBF1, LeCBF1 and TaCBF1 (Jaglo et al., 2001), BnCBF5, BnCBF7 and BnCBF16 (Gao et al., 2002), LeCBF2 and LeCBF3 (Zhang et al., 2004), OsDREB1A and OsDREB1B (Dubouzet et al., 2003), ZmDREB2 (GenBank accession AF450481), HvCBF1 (Xue, 2002), HvCBF2 (Xue, 2003). Unpublished expressed sequence tags (ESTs) and “tentative consensus” sequences (TCs; The Institute for Genomic Research TIGR;
www.tigr.org) were also used for alignments. This list includes sequences from lettuce (LsBQ845452), sunflower (HaBU016783), grape (VvTC35722 and VvTC27620), cotton (GhCA992671), Medicago truncatula (MtTC79347 and MtTC80802), soybean (GmBI967869 and GmAI988637) and rye (ScAF370728 and ScAF370729).

2.4 RESULTS

Truncation Mutants Indicate that the Trans-activity of CBF1 Requires C-terminal to Threonine 158.

Preliminary evidence indicated that CBF1 amino acid residues 115-213 functioned as a transcriptional activation domain (E.J.S and M.F.T., unpublished data). To more precisely define the regions contributing to trans-activity, we introduced four stop codons at threonine 116 (T116), T158, T178 and serine 196 (S196) (Figure 2.1). These truncation mutants were transformed into yeast reporter strain Y190. β−galactosidase assays indicated that truncation mutants T116 and T158 have no enzymatic activity, whereas truncation mutants T178 and S196 retained about 28% and 123% WT (wild type) (GAL4DBD/CBF1_AD) activity level, respectively (Figure 2.2). Thus, these data suggest that the trans-activity requires C-terminal to T158.

Multiple Regions in Arabidopsis CBF1_AD Contribute to Transcriptional Activation.

Concurrent with the generation of the truncation mutants, we also systematically mutated adjacent three to four amino acid residues into alanine. These mutants were
created in the context of \text{GAL4}_{DBD}/\text{CBF1}_{AD} fusions. \(\beta\)-galactosidase activity indicated that alanine scanning mutants between TCK116 and SEG162 were not significantly different from WT \text{GAL4}_{DBD}/\text{CBF1}_{AD} (Figure 2.2). Starting with mutant construct FYM166 and continuing to the C-terminus of \text{CBF1}_{AD}, several mutants were significantly reduced in trans-activity, including FYM166, TMF172, GMP175, TLL178, WNH196, DEE169, EG184 (Figure 2.2). Mutant FYM166 retained about 56% WT activity; TMF172, 59%; GMP175, 77%; TLL178, 74%; WNH196, 63%; DEE169, 82%; and EG184, 75%. The former four mutants were alanine substitutions in either bulky hydrophobic amino acids or aromatic amino acids. The latter two mutants DEE169 and EG184 were substitutions in acidic amino acid residues (Figure 2.2). Mutation in the last three amino acids WSY resulted in a significant increase of trans-activation. Thus, results from \(\beta\)-galactosidase activity indicate that (1) there are multiple regions containing hydrophobic and aromatic amino acids from C-terminal end to F166 contribute significantly to trans-activation; (2) acidic amino acids play a minor role in trans-activation; (3) no single amino acid is critical for trans-activation; (4) the last three amino acids have negative effect on trans-activity.

**Hydrophobic Cluster Analysis, a Computational Approach to Predict Hydrophobic Motifs**

Alanine scanning mutagenesis suggested that no single amino acid was responsible for trans-activation and those mutants with reduced trans-activity contained aromatic or bulky hydrophobic amino acids. These data raised the possibility that multiple
aromatic/hydrophobic groups contributed towards trans-activation. To identify those residues and groups, we reasoned that we might need to mutate multiple amino acids into alanines. To identify candidate residues to alter, we used HCA (Hydrophobic Cluster Analysis).

Earlier studies of the Varicella-Zoster virus trans-activation protein ORF10 program utilized a program HCA (Lemesle-Varloot et al., 1992; Callebaut et al., 1997), to predict key residues in trans-activation (Moriuchi et al., 1995). HCA of CBF1\textsubscript{AD} identified six hydrophobic clusters (HC) (Figures 2.1 and 2.2). Nonsense codon mutants and alanine scanning mutants indicated that trans-activation activity resided C-terminal to T158. As such, we did not focus on HC1 (LMMVIY146). Each of the additional five clusters, HC2 to HC6, included hydrophobic and aromatic amino acids whose substitution with alanine resulted in reduced trans-activity (Figures 2.2 and 2.3). HC2 (FYM166) was included in the first round alanine scanning mutagenesis and retained about 56% WT activity (Figure 2.2). Residues in HC3 (M173F174M176) had been tested in two mutants, TMF172 and GMP175, with 60% and 77% of WT activity respectively (Figures 2.1, 2.2, and 2.3). Residues in HC4 (L179 L180 M183 M187 L188 L189) had been included into three mutants, TLL178, DNM181, and MLL187. TLL178 retained about 74% WT activity, whereas the last two mutants had about the same level of WT activity (Figures 2.1, 2.2, and 2.3).

Two additional HCs, HC5 and HC6, were separated from the other four (HC1-4) HCs by three prolines (Figures 2.1, 2.2). HC5 (V194 Q195 W196 Y200) is a mosaic cluster, in which regular alternations of hydrophobic and non-hydrophobic residues occur.
HC5 residues span three mutants, SVQ193, WNH196 and NYD199, whereas the WNH196 resulted in one of the greatest reductions (about 63% WT activity) in trans-activation in the first round alanine substitution analyses (Figures 2.1, 2.2). The other two, SVQ193 and NYD199, had no obvious effect to reduce trans-activation (SVQ193, 124%; for NYD199, 129%) (Figure 2.2). To determine which amino acid residue in HC5 was important for integrity of the cluster, we substituted each amino acid residue in HC5 with alanine and then performed HCA analyses. This \textit{in silico} experiment indicated that substitution of W196 with alanine results in the loss of HC5 integrity, whereas the substitutions of other residues (V194, Q195 and Y200) did not have this effect. Furthermore, change of S193 into alanine resulted in about 111% WT activity (Data not shown). These results suggest that W196 is more important than other residues in HC5 for trans-activation, which is consistent with results from the mutant of WNH196. HC6 consisted of V208, L210, W211 and Y213, at the C-terminal end of CBF1\textsubscript{AD} (Figures 2.1, 2.2, and 2.3). These residues were separated into two mutants, VSL208 and WSY211. The VSL208AAA was not significant different from WT, whereas the WSY211 mutant resulted in a two fold increase over WT activity (Figure 2.2). Collectively, HCA analyses resulted in prediction of several hydrophobic clusters composed of aromatic and hydrophobic amino acid residues.

\textbf{Multiple Hydrophobic Cluster Motifs Redundantly Contribute to Trans-activation.} Alanine scanning mutants and HCA computational modeling together suggested that there were multiple aromatic and bulky hydrophobic motifs contributing to trans-
activation. Based on these analyses, we hypothesized that there were five hydrophobic motifs within the CBF1\textsubscript{AD} that contributed toward the total trans-activation activity. To test this hypothesis, we changed all amino acid residues in every HC and then combined HC mutants into single constructs to test the effect on trans-activation. Results from these experiments showed that independent elimination of HC3 or HC4 reduced activity to 74% and 70% WT activity respectively (Figure 2.4). One variant, alanine substitutions of HC3 + L179 (one residue of HC4), decreased its activity slightly more, to 50% WT (Figure 2.4). Another variant, alanine substitution of HC4 except the L179 (HC4 – L179), had slightly higher trans-activation activity of 84% WT compared to HC4 of 70% WT activity (Figure 2.4). These data suggest that individual residues may function in an additive way.

Combining the HC3 and HC4 mutations into a single construct resulted in trans-activity that was 25% of WT (Figure 2.4). The triple mutant, HC2/HC3/HC4, decreased the β–galactosidase level to about 7% of WT (Figure 2.4). Based on these analyses, we conclude that Arabidopsis CBF1\textsubscript{AD} harbors multiple hydrophobic motifs that redundantly contribute toward its trans-activity.

**HCs are also Critical for Trans-activation in Plants**

To determine whether the results obtained using yeast were indicative of plants, we then overexpressed these mutants in a GAL4 responsive β- guluronidase Arabidopsis reporter construct.
RNA gel blot indicated that GAL4\textsubscript{DBD}/CBF1\textsubscript{AD} (WT) activated the *uidA* reporter gene (Figure 2.5; lane 1 and 2), whereas GAL4\textsubscript{DBD} did not (Figure 2.5; lane 3, 4 and 5). Alanine substitutions in HC3 or HC4 had no significant effect on reporter gene level (Figure 2.5; lanes 10-13 and lane 19 respectively). Similarly, mutants HC3 + L179 and HC4 – L179, noticeably did not alter reporter gene expression levels. Combining HC3 and HC4 mutations into a single construct significantly reduced the reporter gene expression level (Figure 2.5; lanes 16-18). The triple mutant, HC2/HC3/HC4, was completely below the detectable level (Figure 2.5; lanes 6-9). In all lines, the GAL4 fusion was expressed at relatively similar level (Figure 2.5 middle panel) when normalized to the *eIF-4A* control (Figure 2.5 lower panel).

Based on these experiments, we conclude that alanine substitutions of only one HC (HC3 or HC4) is not sufficient to reduce trans-activation in plants; disruption of two or more HCs are required to reduce the trans-activation; substitutions of all residues within three HCs (HC2/HC3/HC4) are able to eliminate the trans-activity in plants to an undetectable level by RNA gel blot analysis.

**The Hydrophobic Clusters Important for the Activating Capacity in Arabidopsis CBF1\textsubscript{AD} are Conserved across Plant Taxa.**

Alignment of AtCBF1, AtCBF2 and AtCBF3 indicated that these hydrophobic cluster patterns were conserved among three Arabidopsis CBFs (Figure 2.1). To determine whether this pattern is conserved among the CBF proteins from other plant species, we searched the database for AP2 domain containing protein sequences with two
CBF signatures and identified 30 homologous sequences. The C-terminus of these sequences was then aligned using Clustal X. The alignment revealed that the hydrophobic cluster patterns were conserved among other plant species, both in monocot and in dicot (Figure 2.6 A and B). HC2 was separated from HC3 by three acidic amino acids. HC3 and HC4 were separated by a proline in dicots, but form a large cluster in monocots. Two or three conserved prolines separate HC4 from HC5. HC6 is the best-conserved cluster residing at the end and separated from HC5 by a significant number of acidic amino acids and other small amino acid residues (Figure 2.6A and B).

Interestingly, two low temperature sensitive crop species, tomato and soybean, are less conserved in at least one HC. Tomato homologous CBFs are less conserved in HC2 and HC5, whereas soybean CBFs are less conserved in HC3 and HC5 (Figure 2.6A).

There were also subtle differences between dicot and monocot CBFs. In dicots, HC2 and HC3 are separated by three or more acidic residues. In monocots, these acidic residues are dispersed in each HC cluster (Figure 2.6 B). Monocot sequences also lack the conserved proline residue that separates HC3 from HC4 in dicots. This absence results in a longer and contiguous HC3/HC4 cluster. Monocot CBFs also tended to have a Trp residue in HC3 and Tyr in HC4 instead of Leu (Figure 2.6B). Several monocot and dicot CBFs had one proline on the NH3 side of HC6 (Figure 2.6B).

From above analyses, we conclude that plant CBF activation domains are a highly ordered structure. Each hydrophobic cluster motif identified in mutational analyses is also conserved among plant species. Furthermore, chemically conserved amino acid residues divide these stretches of aromatic and/or hydrophobic amino acids into different
motifs. Subtle motif difference in cold sensitive and cold tolerant species may imply their functional difference in low temperature tolerance.

**Growth Stunting is Related to the Potency of an Activation Domain.**

During the screening of transgenic seedlings overexpressing the GAL4$_{DBD}$/CBF1$_{AD}$ construct, we observed that these plants exhibited a stunted growth phenotype. In contrast, the GAL4$_{DBD}$/CBF1$_{AD}$ triple HC mutant constructs were phenotypically normal (Figure 2.7).

Constitutive overexpression of full length CBFs also resulted in growth stunting (Gilmour et al., 2000). This phenotype is thought to result from constitutive expression of the downstream COR genes under normal growth condition. However, overexpression of GAL4$_{DBD}$/CBF1$_{AD}$ in plants should not activate COR gene expression, because this fusion protein lacks the CBF DNA binding domain (not shown). These data of GAL4$_{DBD}$/CBF1$_{AD}$ overexpression in plants suggest that overexpressed activation domain may explain the molecular mechanism for the growth stunting phenotype.

### 2.5 DISCUSSION

Arabidopsis CBF1 plays a major role in plant cold acclimation. Interestingly, many cold sensitive crop species like banana also contain homologous CBFs like these cold tolerant species. Overexpression of Arabidopsis CBF1 in Arabidopsis activates COR genes without low temperature treatment (Jaglo-Ottosen et al., 1998). This suggests that
CBF contains all necessary features to ensure activation. To understand the molecular mechanism for trans-activation of plant cold-regulated genes, we used a biochemical strategy. Our extensive mutagenesis revealed that Arabidopsis contains multiple hydrophobic cluster motifs, compared to two motifs in human herpes simplex virion protein VP16 (Regier et al., 1993) and one in maize C1 (Sainz et al., 1997). These motifs redundantly contribute to trans-activation, providing advantages to protect plant from extreme stresses or mutations. Extra copies might guarantee the necessary activation of stress-related genes when needed.

Multiple hydrophobic motifs have the advantage to recruit different transcriptional cofactors simultaneously to the promoter region. Such a strategy would be useful to activate gene expression quickly in respond to environmental stimuli. Yeast acidic transcriptional activator Gcn4 contains seven HCs (Jackson et al., 1996). Studies of Gcn4 indicate that the binding to coactivators TFIID and Ada-Gcn5 complexes, is dependent on the hydrophobic cluster motifs (Drysdale et al., 1998). Other work in yeast with transcription factor Hap4 indicate that its two transactivation regions use different cofactors, either Gcn5 dependent or Gcn5 independent (Stebbins and Triezenberg, 2004). Investigation on VP16 indicates that large-scale chromatin decondensation is a common effect of acidic activation domain (Carpenter et al., 2005). The transcriptional activation activity and chromatin decondensation has been shown to be separable by using short acid hydrophobic peptides (Carpenter et al., 2005). Furthermore, VP16 acidic activation domain (AAD) act its the large-scale chromatin remodeling activity in part through the recruitment of histone acetyltransferase complex to promoter region (Kuo et al., 1998).
Research indicates that $\text{VP16}_{\text{AAD}}$ also interacts with several other components. Comparison of different HCs in Arabidopsis acidic $\text{CBF1}_{\text{AD}}$ and $\text{VP16}_{\text{AAD}}$ indicates that these motifs are composed of different residues with varied length. Thus, these differences might be the reason that they could recruit different cofactors to bring to the promoter region.

Arabidopsis CBF1 can also interact with Arabidopsis GCN5 and ADA2 proteins (Stockinger et al., 2001). GCN5 and ADA2 proteins are key components of yeast ADA and SAGA histone acetyltransferase complex. In yeast, CBF1 requires yeast ADA2, ADA3 and GCN5 activity in order to be fully functional (Stockinger et al., 2001). Thus CBF1 might recruit ADA or SAGA like histone acetyltransferase complexes to the promoter region to modify chromatin to facilitate transcriptional activation.

Our mutational analyses demonstrated that the predominant acidic amino acids have a minor contribution toward to the transactivation. Research in maize C1 indicated that change of a aspartate (D262) residue to valine resulted in about one fourth of WT activity (Sainz et al., 1997). Mutation of glutamate 476 in VP16 C-terminal activation region also showed a reduction in trans-activation (Sullivan et al., 1998). Thus, for this long-standing puzzle about the role of predominant acidic residues, it seems that these acidic amino acids provide a minor contribution to transactivation. They possibly act by either making direct, specific contact with its target protein factor or establishing long-range electrostatic interactions with and then attract basic protein factors closer in solution (Uesugi et al., 1997).
Based on our findings and the above discussions, we propose a model to describe our understanding how Arabidopsis CBF1 activates downstream plant cold-regulated genes. After sensing low temperature, Arabidopsis plants induce the expression of CBF1. Presently this is activated through unknown signal pathways. CBF1 protein is delivered into the nucleus and binds to its target sites. After binding, the acidic residues in the CBF1$_{AD}$ play their roles by electrostatic force to attract associated cofactors. The hydrophobic clusters in CBF1$_{AD}$ then undergo a conformation change from a random coil to an $\alpha$–helix during or after interacting its associated protein factors. During this process, different hydrophobic clusters associate with different cofactors having different functional activities at promoter region, such as cofactors involved in histone acetylation or methylation.

Only overexpression of activation domain in Arabidopsis also resulted in stunted growth phenotype providing us interesting insight into their mechanisms. Many transcription factors have the potential to be used in improvement of agricultural crops through biotechnology. However, the dwarf phenotype associated with TF overexpression makes that improvement strategy less attractive, simply due to lower yield caused by stunted growth. Farmers do not want crop cultivars with improved stress tolerance but lower yield. When overexpression of a TF in plants, it usually activates a set of genes constitutively (Jaglo-Ottosen et al., 1998), making interpretation of dwarf phenotype complicated. In our investigation, overexpression of GAL4$_{DBD}$/CBF1$_{AD}$ in Arabidopsis also resulted in stunted growth. Our observations suggest that the severity of
growth retardation is related to the potency of activation domain. Our studies imply that the activation domain is related to the growth retardation phenotype.

2.6 ACKNOWLEDGMENTS

We thank Amanda L. Smith and Daniel P. Mathur who generated many of the alanine substitution mutants. We also thank Dr. John J. Finer, who provided space in his growth room. This research was supported in part by the Department of Horticulture and Crop Science, the Ohio Agricultural Research and Development Center (OARDC) at the Ohio State University and OSU Seed Grant Award SG101136 to E. J. S.; an OARDC Graduate Research Competition Grants to Z. W. Portions of this chapter have been published in Plant Molecular Biology (2005) 58: 543-559.

2.7 REFERENCES


Xue, G.P. (2002). An AP2 domain transcription factor HvCBF1 activates expression of cold-responsive genes in barley through interaction with a (G/a)(C/t)CGAC motif. Biochim Biophys Acta 1577, 63-72.


Figure 2.1 Alignment of the C-terminal 98 amino acid sequence (116-213) of Arabidopsis CBF1 with that of CBF2 and CBF3.

Stop codons were introduced at four positions as indicated by arrows, T116 (bolded red), T158 (bolded pink), T178 (bolded brown), and S193 (bolded turquoise). Residues in bolded italic are bulky hydrophobic or aromatic amino acids composed of six different hydrophobic clusters. Each cluster is indicated by underline. The asterisk (•) is used to indicate the identical amino acid residues between CBF1, CBF2, and/or CBF3. The minus symbol (–) is used to indicate a gap in the alignment.
Figure 2.2 β- galactosidase activities of truncation and alanine substitution mutants.

The text below each bar indicates the construct used for galactosidase enzymatic assays, including WT GAL4/CBF, four truncation mutants with introduction of stop codon at T116, T158, T178 and S193, and 29 additional alanine substitution mutants. Below the bar showing the activity of each mutant, the amino acid residues mutated into alanines are indicated followed by the position number of the first amino acid residue mutated. For example, TCK116 is altered in that the three amino acids Thr, Cys, and Lys, which were mutated into alanine with Thr as the 116th amino acid in CBF1. All constructs were transformed into yeast reporter strain Y190. Five independent transformants from a single transformation event were inoculated into media and used to conduct the assay. These assays were repeated multiple times. Error bar is the standard error of the mean.
Figure 2.3 Hydrophobic Cluster Analysis (HCA) of CBF1 carboxy terminal 101 amino acid residues revealed six hydrophobic cluster (HC) motifs.

HC1 is composed of LMMVIY; HC2, FYM; HC3, MFM; HC4, LLMMLL, HC5, VWY; HC6, VLWY. Arrows also indicate the positions of the four truncation mutants created by the introduction of stop codons.
Figure 2.4 β-galactosidase activities of hydrophobic cluster mutants.

Single, double, triple HC mutants, and two HC derivatives are indicated on top of each bar. The numbers following each HC mutant indicate the position of each HC’s first amino acid residue relative to the full length CBF1. Five independent transformants were inoculated for each assay. Assay was repeated multiple times. Error bar is the standard error of the mean.
Figure 2.5 RNA gel blot analysis indicates that hydrophobic clusters critical for CBF1 activation domain transactivation in yeast are also critical in Arabidopsis.

WT (WT), GAL4DBD, and HC mutants were transformed into an Arabidopsis reporter line harboring a single locus insertion of a UAS$_{G17mer(x4)35S(-46)}$ uidA reporter gene. The different effector or activator constructs used for transformation were listed above lanes. Independent transgenic seedlings were randomly separated into two to four pools and eight μg total RNA from each pool with at least 10 seedlings were hybridized. The same filter was hybridized, stripped, and rehybridized with probe for uidA, GAL4DBD, and eIF4A. eIF-4A was used as loading control.
CBF sequences from dicots (A) or monocots (B) are identified using a two letter prefix code representative of binomial nomenclature (genus and species). Common names for plants: At (Arabidopsis), Bn (canola), Le (tomato), Ls (lettuce), Ha (sunflower), Vv (grape), Gh (cotton), Mt (Medicago truncatula), Gm (soybean), Os (rice), Zm (maize), Hv (barley), Ta (wheat) and Sc (rye). Published sequences are identified according to their published name. Expressed sequence tags (ESTs) and sequences from GenBank database are followed by the GenBank or TC accession number. Different hydrophobic clusters are overlined. Strongly hydrophobic residues V, I, L, F and moderately hydrophobic residues M, W, Y, and A are turquoise; polar noncharged residues C, N, Q, S, and T are green; P is yellow; G is orange; positive charged residues H, K, and R are red; negatively charged residues D and E are white letter on a purple background.
Figure 2.7 Non-activating HC2/3/4 mutants exhibit a normal growth phenotype

Overexpression of WT (GAL4/CBF1) in Arabidopsis reporter line harboring UAS$_{G17mer}$G35S(-46) $uidA$ reporter construct showed growth stunting compared to triple HC mutant.
CHAPTER 3

EXTENSIVE MUTAGENESIS REVEAL A HIGHLY ORDERED STRUCTURE ADOPTED BY ARABIDOPSIS TRANS-ACTIVATOR CBF1 TO FULFILL ITS FUNCTIONAL ROLES

3.1 ABSTRACT

Understanding the molecular mechanism of how plant transcription factors regulate gene expression is of fundamental and agricultural importance. Our previous work indicated that four hydrophobic stretches in the Arabidopsis CBF1 activation domain provided functional redundancy in trans-activation, and this potency of trans-activation was related to the severity of stunted growth phenotype when overexpressing activators in plants. Further analyses suggested that while the hydrophobic residue Met in the GMP175 mutant was indeed responsible for trans-activation, the different hydrophobicity resulting from the mutation might be the final reason for the transcriptional difference. This was also supported by the mutation of the hydrophilic Ser residue in the WSY motif into a hydrophobic Ala residue resulting in much higher trans-activity. The high conservation of the WSY motif with its inhibitive effect on trans-activation suggests an unknown self-control mechanism in the regulation of downstream plant cold-regulated
(COR) genes. Sequence comparison revealed that two alanine stretches were conserved after the AP2 domain and before the activation region of CBF proteins. Furthermore, leucine substitution of the conserved alanine residues reduced the capacity for trans-activation, supporting our hypothesis that the restoration of stunted growth phenotype in transgenic CBF lines was related to the potency of CBF1 trans-activation. This was also revealed by the lower activating activity in GAL4/VP16 lines with less severe stunted growth than GAL4/CBF1 lines. Collectively, our extensive mutagenesis reveals that the conserved alanine residues, hydrophobic stretches, and WSY motif residing before, within, or at the very C-terminal end of the CBF1 activation domain, have specific roles in the regulation of gene expression. Thus, our biochemical analyses of CBF1 provide insight into how a plant transcription factor can adopt a highly ordered structure to modulate gene expression toward meeting its specific roles.

3.2 INTRODUCTION

Plants are sessile. They have to cope with and adapt to changing environments. One strategy to do this is by the continual regulation of gene expression in response to external and internal cues during their life cycles. During these processes, it is critical for transcription factors (TFs) providing specific and coordinated transcriptional modulation of numerous gene expressions. This might be one reason that plants have many more TFs, up to 6% of the genes in the genome, than animals (Arabidopsis Genome Initiative, 2000). Thus, understanding the molecular mechanism of how a plant transcription factor
regulates gene expression is of fundamental and agricultural importance. Despite their importance, the molecular mechanism of how TFs function is less studied in plants compared to other eukaryotes. One central question is how do transcriptional activation domains function. Arabidopsis transcriptional activator CBF1 (C-repeat bind factor 1) and its homologs in crop species play a “master switch” role in the regulation of genes contributing to cold tolerance (Stockinger et al., 1997; Sarhan, 1998). Thus understanding how CBF1 functions as a master switch may provide insight into how plant transcription factors regulate gene expression, and also lead to a greater insight about plant cold acclimation.

The ability to use TFs to activate an entire suite of genes makes the ectopic overexpression of TF in plants to improve agricultural traits an attractive strategy. For example, the overexpression of CBF1 in plants activates many COR genes and significantly improves freezing tolerance, whereas the overexpression of a single COR gene in plants has no observable phenotype at the whole plant level (Artus et al., 1996; Jaglo-Ottosen et al., 1998). However, a negative consequence of the overexpression of TFs in plants is a stunted growth phenotype, thus potentially affecting crop yield. These problems must be overcome by plant biologists because no farmer wants crop cultivars with improved stress tolerance but with lower yields (Zhang et al., 2004). Thus, understanding the molecular mechanism of stunted growth is of particular importance to agricultural biotechnology.

To increase our understanding of the molecular mechanism of transcriptional regulation and its potential link to the stunted growth phenotype, we used mutagenesis.
The effect of CBF1 mutants on trans-activation was then determined by the measurement of the activity of a reporter gene, β–galactosidase. Previously, we found that Arabidopsis CBF1 harbored multiple hydrophobic clusters (HC), or motifs that we referred to as HC1 to HC6. HC2, HC3, HC4, and HC5 redundantly contributed to transactivation (Wang et al., 2005). Nevertheless, an HC2/3/4 triple mutant still retained about 7% WT activity as indicated by β-galactosidase activity. We hypothesized that this residual activity was contributed by HC5.

Our systematic alanine substitutions enabled us to test every residue except alanine residues. To test these alanine residues, we substituted them with leucine and determine the mutation’s effect on trans-activation.

Residues comprising HC6 at the COOH-terminal end of CBF1 are the most highly conserved in CBF activation domains and as such of interest (Wang et al., 2005). HC6 consists of Val-Leu-Trp-Ser-Tyr. Further isolations of more homologous CBFs, about 20 from barley (*Hordeum vulgare*), 10 from rice (*Oryza sativa*) (Skinner et al., 2005), and 11 CBFs from diploid wheat (*Triticum monococcum*) (Miller et al., 2006), confirmed its high conservation. Curiously, in our alanine substitution mutants and in our truncation mutants, the conversion of the last three amino acid residues WSY into alanines resulted in about two-fold increase of β-galactosidase activity (Wang et al., 2005). Here, we then continue our investigation to determine each amino acid residue’s role in transactivation. Also through extensive mutagenesis, we present our understanding that Arabidopsis CBF1 transcriptional activator adopts a highly ordered structure to fine-tune gene expression to meet their specific roles. Conserved stretches of amino acid residues play
specific roles toward the regulation of gene expression. Our analyses provide an example for that how a plant transcription factor adopts a highly ordered structure to modulate gene expression.

3.3 MATERIALS AND METHODS

Site-directed Mutagenesis

Mutagenesis either alanine substitution or leucine substitution was conducted using primers that incorporated mutations at the target amino acids (Li and Wilkinson, 1997) with the QuickChange strategy (Stratagene, La Jolla, CA). All mutations were introduced into a GAL4$_{DBD}$CBF1$_{AD}$ fusion that was in plasmid pGEM11Z- (Promega, Madison, WI). After confirmation by sequencing, we then subcloned the GAL4$_{DBD}$CBF$_{MUTANTS}$ into a yeast expression vector pDB20.1 (Berger et al., 1992), and a plant expression vector, pCIB710 (Rothstein et al., 1987). Constructs in pCIB710 were further subcloned into the Agrobacterium tumefaciens binary vector pCIT20 (Ma et al., 1992).

Yeast β−galactosidase Assays

For β−galactosidase assay, we transformed plasmids expressing GAL4$_{DBD}$CBF1$_{AD}$ (treated as wild type) or GAL4$_{DBD}$CBF1$_{MUTANTS}$ into yeast reporter strain Y190 harboring a reporter construct, which was described previously (Wang et al., 2005). To assay the activity of β-galactosidase produced by the reporter construct, the extracts from the cultures of five independently transformed yeast colonies were used. The measurement
of β-galactosidase activity used the procedure of Wang et al. (2005) and results were normalized to the protein concentration of samples as determined by Bradford assay (Bradford, 1976).

**Plant Transformation and Transgenic Seedling Screening**

Plant expression constructs were electroporated into the *Agrobacterium* strain GV3101 (Koncz, 1986). Arabidopsis plants (ecotype WS-2) were transformed by the floral dip method (Clough and Bent, 1998). Transgenic seedlings were screened on GB5 media supplemented with hygromycin 30 μg/ml.

**RNA Isolation and RNA Gel Blot Analyses**

Independent T1 seedlings screened by selective media were randomly pooled. Each pool contained at least 10 transgenic seedlings. Total RNA was isolated using Plant RNeasy Mini Kits (Qiagen Inc., Valencia, CA). RNA gel blot analyses were conducted as described previously (Wang et al., 2005). Filters were hybridized, stripped, and rehybridized using the following gene probes, *GAL4DBD*, *uidA*, and *eIF4A*.

**Sequence Analyses**

AP2 domain containing sequences from dicots used for alignments were identified previously (Wang et al., 2005). Sequence alignments begin with the COOH-terminal CBF signature, DSAWR. Alignments were conducted using ClustalW ([http://www.ebi.ac.uk/clustalw/](http://www.ebi.ac.uk/clustalw/)) (Thompson et al., 1994). Aligned sequences were formatted for viewing using BOXSHADE 3.21 ([www.ch.embnet.org/software/ BOX form.html](http://www.ch.embnet.org/software/BOX form.html)). The prediction of a potential phosphorylation site within CBF1 was conducted using NetPho ([http://www.cbs.dtu.dk/services/NetPhos/](http://www.cbs.dtu.dk/services/NetPhos/)) (Blom et al., 1999).
3.4 RESULTS

Activity of WNH196 Mutant in Plants is Not Significantly Different from Wild Type

Previously, the HC2, HC3, and HC4 of CBF1 activation domain were found to be important in plants (Wang et al., 2005). Alanine substitution through these three clusters (triple mutant HC2/3/4) eliminated reporter gene activation in plants (Wang et al., 2005). However, the HC2/3/4 triple mutant retained about 7% residual activity in yeast. We hypothesized that the residual activity of triple mutant HC2/3/4 was contributed by HC5, because the WNH196 mutant was about 65% wild type (WT). HC5 encompassed residues V194 through Y200. Within these residues, W196 appears to be important for the integrity of HC5, as the substitution of this residue with alanine abolishes HC5 when analyzed by Hydrophobic Cluster Analysis (Wang et al., 2005). To test whether the WNH196 mutant was significantly affected in plants, we determined the expression levels of a GAL4 responsive gene in seedlings overexpressing GAL4/CBF1 (WT) and GAL4/CBF1_{WNH196}. The results from this experiment indicated that the WNH196 mutants were not significantly altered in the level of trans-activity compared to that of WT (Figure 3.1).
Single Amino Acid Substitutions through GMP175 Indicates that the Activity of M176A is Reduced whereas the Activity of P177A is Increased

Alanine substitution through the COOH-terminal 98 amino acid residues of the CBF1 activation domain indicated that GMP175 mutant was about 75% WT activity (Wang et al., 2005). To determine which individual residue within the GMP175 mutant was responsible for trans-activation, we substituted each individual residue with alanine. Individual substitution of each residue through GMP indicated that G175A, M176A, and P177A were 112%, 79%, and of 166% WT respectively (Figure 3.2). These data suggest that M176 has a positive role in transactivation and P177 has an inhibitory role.

The Highly Conserved WSY Motif has an Inhibitive Effect on Trans-activation in Yeast and Arabidopsis

Comparative sequence alignments of CBF from many plants reveal that the C-terminus is highly conserved (Wang et al., 2005). In dicots there is a stretch of six residues consisting of VSLWSY (Figure 3.3), and in monocots there is a four-residue stretch consisting of LWSY present in virtually every CBF homolog identified (Chapter2, figure 2.6) (Wang et al. 2005).

Previously, we observed that the WSY211 mutant resulted in about two-fold increased activity over WT when assayed in yeast (Wang et al., 2005). To determine which residue was responsible for this increased activity, we converted each individual amino acid residue in the WSY motif into alanine, and then assayed the trans-activity of mutants in a yeast reporter strain. Surprisingly, all three alanine-substitution mutants
resulted in about two-fold increased activity over WT. W211A mutant had about 167% WT activity; S212A, 209%; Y213A, 202% (Figure 3.4A).

To determine whether these mutants would be similarly affected in Arabidopsis, we overexpressed them in the Arabidopsis reporter line (Wang et al., 2005). RNA gel blot analysis indicated that uidA reporter gene expression levels were higher than WT for two of these mutants (Figure 3.4 B and C). S212A was 159% WT activity and Y213A was 147% WT activity (Figure 3.4 B and C). W211A was no different from WT. Thus, S212A had the highest activity in plants and yeast.

**Leucine Substitution of Alanine Residues Reduced Trans-activity**

Comparative sequence alignment of CBF proteins from many plants revealed that there were numerous highly conserved alanine residues (Figure 3.3). There were two stretches, AAA124-126 and AAA128-131 (Figure 3.3). To determine whether residue size was important for activity, we substituted the alanine residues (small) with leucine, a large bulky hydrophobic amino acid.

We then overexpressed leucine mutants A118L, AAA124LLL, AAA128LLL, and A165L in Arabidopsis reporter line. Results from this experiment revealed that three of the four leucine substitution mutants were reduced in their trans-activity (Figure 3.5). The AAA124LLL was 63% WT activity; AAA128LLL was 74%; and A165L was 66%. A118L was not significantly different from WT and was also the one alanine that was not well-conserved across the different plant taxa (Figure 3.3). These data suggest that the
conserved alanine residues may play an important functional role for the maximal transcriptional activity.

The Activity of the Arabidopsis CBF1 Activation Domain is Greater than the VP16 Activation Domain

VP16, herpes simplex virus VP16, is the prototypical mammalian transcriptional activator. It is one of the most well-studied and most potent activation domains (Sadowski et al., 1988). VP16 activation region is widely used in plant biology for instance in the GAL4/VP16 fusion used in the “enhancer-trap” strategy (Laplaze et al., 2005). To compare the trans-activity of VP16 with CBF1, we overexpressed GAL4/VP16 and GAL4/CBF1 in Arabidopsis. RNA gel blot analyses revealed that uidA expression levels were more than two fold higher in the GAL4/CBF1 overexpressing lines than in the GAL4/VP16 overexpressing lines (Figure 3.6). Thus we conclude that CBF1 activation domain has twice the activation capacity as the potent VP16 activation domain when fused to the GAL4DBD in plants.

3.5 DISCUSSION

Our previous mutagenesis revealed that there were at least three HCs (HC2, HC3, and HC4), comprised of hydrophobic and aromatic amino acid residues, redundantly contributing to trans-activation. The triple mutant (HC2/3/4) eliminated more than 90% activity when assayed in yeast, and no reporter gene expression was detected by this
same mutant in Arabidopsis. In yeast, the WNH196 mutation (part of HC5) reduced the CBF1 trans-activity to about 65% WT (Wang et al., 2005). However, our RNA gel blot results with WNH196 mutant in Arabidopsis did not show obvious difference from WT. The difference between yeast and Arabidopsis data might result from less sensitivity using RNA gel blot than the β-galactosidase assay (Figure 3.1).

Comparative sequence alignments indicated that the WSY motif was the most conserved region in CBF sequences from either monocot or dicot crop species (Figure 3.3) (Skinner et al., 2005; Wang et al., 2005; Miller et al., 2006; Zhang, 2006). Interestingly, the mutations of all three amino acid residues or a single residue of WSY into alanine resulted in almost two-fold increased activity over WT when assayed in yeast (Figure 3.4). This is also supported by our results of truncation mutant S193 STOP. The GAL4 fusion protein without the C-terminal end 20 amino acids of CBF1_{AD} (no HC5 and HC6), trans-activity was higher than WT (Wang et al., 2005). In Arabidopsis, RNA gel blot analyses revealed a similar result with increased trans-activity after the mutation of WSY residues to alanines (Figure 3.4). Curiously, the mutation of S212 residue (S212A), rather than the aromatic residues W211 and Y213 resulted in a big increase in trans-activity both in yeast and in Arabidopsis. Serine is a small hydrophilic amino acid, and tends to interact with the aqueous environment to form H-bonds. Ser residues are predominantly found on the protein exterior surfaces. Detailed comparison of published mutational analysis data from VP16 and GCN4 revealed a commonality in that the mutants of VP16 or GCN4 with lowest trans-activity contained a Ser substitution (Jackson et al., 1996; Sullivan et al., 1998). For example, double mutant F479S D486G
had 10% WT activity; triple mutant F475S M478T F479S, <1%; double mutant F473S E474G, 12%; F473S, 26% (Sullivan et al., 1998). Sullivan et al. (1998) also substituted F473, F475, M478, and F479 of VP16 to 3 to 13 different amino acid residues. Among these mutants, almost all Ser substitution mutants had the lowest activity. A similar finding was also made for GCN4 (Jackson et al., 1996). In CBF1 there are no Ser or Thr residues present as part of hydrophobic clusters (HC2, FYM; HC3, MFGM; and HC4, LLDNMAEGMLL). Ser and Thr are only present between these stretches. There is one Thr residing in HC1 (LDMEETMVEAIY), however HC1 contributes no trans-activity (Wang et al., 2005). There are however two Ser residues in HC6, VSLWSY (Figure 3.3). Intriguingly, LeCBF1 and LeCBF2 also contain a Ser residue COOH terminal to WNY. No Ser or Thr is present in the VP16 activating motif either (Sullivan et al., 1998).

Because it is a common mechanism to regulate the TF’s function by phosphorylation (Whitmarsh and Davis, 2000), and the phosphorylation often occurs on the Ser residue, we then ask whether the VSLWSY is a potential phosphorylation site and phosphorylation plays a role in the trans-activity. NetPho program with a sensitivity, the proportion of true phosphorylation sites predicted, ranging from 69% to 96% was used (Blom et al., 1999). Prediction results indicated that S209 had a score of 0.038; S212, 0.044 (Figure 3.7). Either S209 or S212 is much lower than the threshold score 0.5, which suggest that phosphorylation might not play a role in this inhibitory effect of WSY motif. We think it is possible that the hydrophilicity from two Ser residues within VSLWSY motif might be inhibitory for trans-activation. The conserved polar acidic amino acid residue stretch (DGEGE in CBF1) immediately NH3-terminal to the
VSLWSY motif may also help HC6 to swing outside of the protein surface. We propose a self-controlled mechanism that CBF1 applies this negative effect on trans-activation to fine-tune its downstream plant COR gene expression during plant cold response. In this model, HC6 with bulky aromatic residues Trp (W) and Tyr (Y) flanking S212 can cause HC6 to swing like a horse tail onto the surface of the modular activation domain (HC2-HC5) to block or reduce the chance of protein-protein interaction between positive HCs, HC2-5, with its associated transcriptional cofactors. This protein-protein interaction is critical for transcriptional activation. I call this the “MaWeiBa Gongnong” (Chinese translation for the function of a horse tail) model to illustrate CBF1’s fine-tuning the expression of downstream genes (Figure 3.8).

Key residues within hydrophobic stretches affect regional hydrophobicity and hydrophilicity can affect its trans-activity. One previous experiment suggested a relationship between hydrophobicity of hydrophobic clusters and transactivation activity (Almlof et al., 1997). Several identified hydrophobic residues from human glucocorticoid receptor (GR) are critical for trans-activation prompted these authors to propose a relationship between hydrophobicity and trans-activity. For example, I193 of GR was mutated to a residue F with higher hydrophobicity to result in about 150% WT activity. Our mutagenesis of GMP175 individual residue also suggests that the hydrophobicity plays a role in the determination of trans-activity (Figure 3.2). Met is not an aromatic residue with aromatic rings but a hydrophobic residue with higher hydrophobicity than Ala according to the hydrophobicity scale F, I>V>L, M, W>Y>A>G>S>P (http://prowl.rockefeller.edu/aainfo/contents.htm) (Rose et al., 1985).
The P177 residue (and only this one) in CBF1 activation domain separates two clusters, HC3 and HC4. Besides Ala has a higher hydrophobicity than Pro, the P177A substitution resulted in an extended hydrophobic cluster, comprised of residues from HC3 and HC4 (MFMLLMMLL173) (Figure 3.3). This presumably increased hydrophobicity might contribute to the dramatic increase of trans-activation, about 177% WT activity (Figure 3.2). Similarly, the G175A mutation resulted in a moderate increase of hydrophobicity, 120% WT activity. The M176A lost the critical hydrophobic residue Met and decreased hydrophobicity, resulting 85% WT activity (Figure 3.2). The GMP175AAA lost the Met residue, resulting a less hydrophobicity. Notice that the mutation of GMP175 into three alanine residues will not be able to extend the HC, because then four residues (AAAT) would separate HC3 and HC4. Intriguingly, the residue P175, which is highly conserved in dicot CBFs to separate HC3 from HC4, is missing in monocot CBFs (Wang et al., 2005). Thus, our data and published data support this hypothesis that the hydrophobicity of a residue is important in the determination of trans-activity (Almlof et al., 1997; Sullivan et al., 1998; Wang et al., 2005). We summarize our analyses about the relationship between the hydrophobicity or hydrophilicity of an amino acid residue stretch and the resulting activating activity (Table 3.1).

Transgenic seedlings overexpressing transcription factors have the ability to enhance agricultural traits significantly, because they can activate multiple downstream genes (Jaglo-Ottosen et al., 1998). However, the simultaneous growth stunting with potential loss of yields present a challenge to the use of TFs (Zhang et al., 2004). Using a stress-inducible promoter has been proposed as a means to drive the expression of
transcriptional activators to overcome these problems (Kasuga et al., 1999; Lee, 2003). Understanding the molecular mechanism is critical to broaden the application of transcription factors in agricultural biotechnology. Our observation suggests that the severity of growth stunting is related to the potency of activation domain (Wang et al., 2005). Transgenic seedlings overexpressing CBF1 hydrophobic cluster mutants with less potency in transactivation developed more normally. Similarly, Arabidopsis L2 line with higher transcript level of overexpressed CBF4 showed a more severe stunted growth phenotype than line L10. In contrast, line L10, which had lower CBF4 expression level, developed similar to untransformed plants (Haake et al., 2002). This was also found to be true for the overexpression of ABF4 (ABA-responsive element binding factor 4) in Arabidopsis. Line A406 with more severe stunted growth phenotype had higher transcript level of ABF4 than line A402 (Kang et al., 2002). Presumably this higher level of transcripts of CBF4 or ABF4 means higher protein levels (Haake et al., 2002). This hypothesis is also supported by our results presented here between comparisons of two different activators. VP16 activation domain is less potent than the CBF1 activation domain (Figure 3.6), and the transgenic seedlings overexpressing VP16\(_{AD}\) are less severe in growth stunting than the seedlings overexpressing the CBF1\(_{AD}\).

Furthermore, Arabidopsis lines overexpressing CBF3 show a stunted growth habit (Gilmour et al., 2000). Mutagenesis of the stunted plants resulted in the restoration of the normal growth habit (M. F. Thomashow, personal communication). Genetic and molecular characterization of these plants revealed that the mutation was in the CBF transgene and occurred in the alanine stretch (M. F. Thomashow, personal...
communication) that was not subjected to mutational analysis in our original study. In the context of GAL4DBD/CBF1AD fusion, we determined that the leucine substitution of several alanine residues from 116 to 213 indeed affect the trans-activity of uidA reporter gene (Figure 3.5), with the lowest activity mutated in alanine stretches A124-126. These alanine stretches are NH3-terminal to the activation region. Thus, through alanine-scanning mutation, truncation mutation, and leucine substitution, our extensive mutagenesis reveals that the conserved different amino acid residue stretches in CBF1 have their specific functions. So far, although it is important to regulate gene expression in plants, only a few plant transcription factors have been fully or partially mutagenized (Sainz et al., 1997; Wang et al., 2005; Krizek and Sulli, 2006). Thus, our biochemically comprehensive analyses of Arabidopsis CBF1 provide an example to understand how a plant transcription factor adopting a highly ordered structure to modulate gene expression.

3.6 ACKNOWLEDGMENTS

We thank Dr. John J. Finer for the use of his growth room. This research was supported in part by the Department of Horticulture and Crop Science, the Ohio Agricultural Research and Development Center (OARDC) at the Ohio State University and OSU Seed Grant Award SG101136 to E. J. S.; an OARDC Graduate Research Competition Grants to Z. W.
3.6 REFERENCES


Table 3.1 Summary of hydrophobic clusters with transcriptional activities.

<table>
<thead>
<tr>
<th>TFs in paper</th>
<th>HC Name in paper</th>
<th>Sequence</th>
<th>Trans-activity</th>
<th>Mutational analyses (% WT activity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBF1</td>
<td>HC1 146 LDMEETMVEA/Y</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HC2 165 FYM</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HC3 173 MPFM</td>
<td>+</td>
<td>M176A 79%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HC4 179 LDNMAEGMLL</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HC5 194 VQWNHNY</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HC6 208 VS LWGY</td>
<td>-</td>
<td>W211A 167%; S212A 209; Y213A 202%</td>
<td></td>
</tr>
<tr>
<td>GR</td>
<td>HR I 191 EDILQDLEE</td>
<td></td>
<td>F191I 38%; F191V 48%; F191L 57%; F191A 44%</td>
<td>l193F 151%; l193L 42%; l193A 32%; L194V 23%; L194A 19%</td>
</tr>
<tr>
<td>VP16</td>
<td>VP16N 439 LDDEELEMMLGDV</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VP16C 470 MADFEFEQME</td>
<td>+</td>
<td>F473S 26%; F473S,E474G 12%; F475S,M478T,F479S &lt;1%; F479S,D486G 10% Compared to F479L 90%; F457L,F479L 100%;</td>
<td></td>
</tr>
</tbody>
</table>

Note the source of published data, CBF1 (Wang et al., 2005), GR (Almlof et al., 1997), VP16N and VP16C (Regier et al., 1993; Sullivan et al., 1998). Residues in italic form a hydrophobic cluster (HC).
Figure 3.1 The uidA expression levels induced by WNH196 and WT.

(A) WT GAL4/CBF and the WNH196 mutant were transformed into Arabidopsis reporter line harboring a single locus insertion of a UAS_{G17mer(x4)35S(-46)} uidA reporter gene. Independent T_1 transgenic seedlings were randomly grouped into three pools. Each pool consists of at least 10 seedlings. Eight µg total RNA from each pool were loaded onto 1.2% agarose gel. The same filter was hybridized, stripped, and rehybridized with GAL4, uidA, and eIF-4A.

(B) Quantitative representation of relative expression of uidA reporter gene (uidA/eIF-4A). The error bars are the standard deviation of the mean.
Figure 3.2 β-galactosidase levels of individual GMP alanine substitution mutants.

The text below each bar identifies each construct used for assay. Constructs include WT GAL4/CBF, GMP175AAA, and the single residue mutants G175A, M176A, and P177A. Each construct was transformed into yeast reporter strain Y190. Five independent yeast colonies were assayed. Assay was repeated three times. The error bar is the standard deviation of the mean.
Figure 3.3 Alignment of the COOH-terminal regions of dicot CBF proteins.

Arabidopsis CBF1 residues 101-213 are shown on the top lines of the alignments. Alanine residues (A118, A124-126, A128-131, and A165) are overlined above the sequence. The second CBF signature, DSAWR, and four hydrophobic clusters (HC2-5) are overlined. Highly conserved WSY motifs and the conserved acidic amino acid residue stretches prior to WSY motifs are also overlined. CBF sequences are identified using two letter prefix code standing for binomial nomenclature (genus and species). Common names for plants: At (Arabidopsis), Bn (canola), Le (tomato), Ls (lettuce), Ha (sunflower), Vv (grape), Gh (cotton), Mt (Medicago truncatula), Gm (soybean).

Expressed sequence tags (ESTs) and homologous sequences from GenBank are identified using a two letter prefix binomial nomenclature code followed by either GenBank or TC accession number. White lettering on black background indicates sequence identity, and black lettering on a grey background indicates sequence similarity.
Figure 3.4 β-galactosidase activity levels, and *uidA* expression levels of WSY mutants.

(A) β-galactosidase activity levels in yeast. The text below each bar identifies each construct used for assay. Each construct was transformed into yeast reporter strain Y190. Five independent yeast colonies were assayed. Assay was repeated three times. The error bar is the standard deviation of the mean.

(B) RNA gel blot analyses of *uidA* Arabidopsis reporter line transformed with WSY mutants. WT GAL4/CBF and alanine substitution of individual WSY residue were transformed into Arabidopsis reporter line harboring a single locus insertion of a UASG17mer(x4)35S(-46)*uidA* reporter gene. Constructs used for transformation are listed above lanes. Independent T₁ transgenic seedlings were randomly grouped into two or three pools depending on the number of available seedlings. Each pool consists of at least 10 seedlings. The same filter was hybridized, stripped, and rehybridized with *GAL4*, *uidA*, and *eIF-4A*.

(C) Quantitative representation of relative expression of *uidA* reporter gene (*uidA/eIF-4A*) shown in (B). The error bars are the standard deviation of the mean.
Figure 3.5 The *uidA* expression levels of leucine scan mutants in Arabidopsis.

(A) WT GAL4/CBF and alanine substitution of alanine residues in CBF1 activation domain were transformed into Arabidopsis reporter line harboring a single locus insertion of a UAS<sub>G17mer(x4)35S(-46)</sub> *uidA* reporter gene. Constructs used for transformation are listed above lanes. Independent T<sub>1</sub> transgenic seedlings were randomly grouped into one, two, or three pools depending on the number of available seedlings. Each pool consists of at least 10 seedlings. The same filter was hybridized, stripped, and rehybridized with *GAL4*, *uidA*, and *eIF4A*.

(B) Quantitative representation of relative expression of *uidA* reporter gene (*uidA/eIF-4A*). The error bars are the standard deviation of the mean.
Figure 3.6 Relative trans-activation activity of the VP16 and CBF1 activation domains.

(A) RNA gel blot analyses of *uidA* reporter gene expression levels. GAL4/CBF, GAL4/VP16, and GAL4 DNA binding domain lacking an activation region were transformed into Arabidopsis reporter line harboring a single locus insertion of a UASG17mer(x4)3SS(-46) *uidA* reporter gene. Independent T1 transgenic seedlings were randomly grouped into three pools with at least 10 seedlings in each pool. The same filter was hybridized, stripped, and rehybridized with GAL4, *uidA*, and *eIF4A*.

(B) Quantitative representation of relative expression of *uidA* reporter gene (*uidA/eIF-4A*) shown in (A). The error bars are the standard deviation of the mean.
Figure 3.7 Predicted phosphorylation sites within CBF1.

The numbers below the X axis indicate the amino acid position. The threshold score is 0.5. Amino acid residue with high score (>0.5) is more likely to be a phosphorylation site.
Figure 3.8 “MaWeiBa Gongnong” model for Arabidopsis trans-activator CBF1, adopting a highly ordered structure to fine-tune its downstream gene expression.

MaWeiBa Gongnong is the Chinese translation for the function of a horsetail. Arabidopsis CBF1 transcriptional activator harbors a N-terminal DNA binding domain (hatched circle) that binds to the CRT/DRE (gray rectangle) and a C-terminal activation domain. These two domains are linked by peptide containing alanine stretches with a role in relieving stunted growth. Within activation domain, four hydrophobic motifs (HC2-5) undergo a conformation change from random coil to α helix (piled black ovals) upon binding to cofactors. The top one (HC5) is separated by a conserved stretches of Pro residues (empty rectangle), resulting in helix bending. Four HCs redundantly contribute toward trans-activation by interacting with different transcriptional cofactors, like chromatin-remodeling factors. The highly conserved WSY (W and Y residues, red stop sign; S residue, oval between them) are linked by a conserved stretches of polar amino acids composed of acidic amino acids, which possibly help WSY motif swinging outside of activation domain like a MaWeiBa (horsetail in Chinese) to negatively modulate downstream gene expression. Ser residues make the motif on the protein surfaces and turn it to be repressive motif by bulky aromatic residues W and Y interfering the interaction of hydrophobic clusters with transcription cofactors, which provide a self-control mechanism to fine-tune gene expression.
CHAPTER 4

CROSSTALK BETWEEN PATHWAYS INVOLVED IN LOW TEMPERATURE SIGNAL TRANSDUCTION

4.1 ABSTRACT

It has become clear that different signal transduction pathways in response to environmental stresses form a complex web to regulate gene expression, as this is suggested by multiple cis-acting elements in a typical plant cold-regulated (COR) gene COR15a. Here, we set up experiments to determine the potential crosstalk between ERFs (ethylene-responsive element binding factor), CBFs (C-repeat binding factor), and ABFs (ABA-responsive element binding factor) mediated signal transduction pathways. We constructed several yeast reporter strains harboring synthetic binding sites or a natural promoter region of COR15a. Our further constitutive phosphorylation mutant of ABF1S104D failed to trans-activate reporter gene in yeasts, suggesting that additional modifications was required. Our results provide useful information in constructing a functional yeast reporter strain.
4.2 INTRODUCTION

In Arabidopsis, it is clear that several different phytohormones including ABA and ethylene play roles in drought and low temperature responsive signal transduction pathways. Pathways mediated by different phytohormones could form a complex web of overlapping signal transduction pathways (Gazzarrini and McCourt, 2003). Ethylene is a simple gaseous hormone that regulates plant growth and development in response to a diverse array of biotic and abiotic signal stimuli. It can be induced by the environmental stresses like flooding or drought (O'Donnell et al., 1996; Chen et al., 2005), which suggests that ethylene plays a role in stress tolerance. The exposure of rye (*Secale cereale*) to ethylene induces antifreeze activity (Yu et al., 2001). However, the molecular mechanism is unknown. The ethylene signal transduction begins from the perception of ethylene by its receptors. The signal is then transmitted via cascades of intermediates. Late in the cascade, ethylene responsive element (ERE) binding factors (EREBP or ERF) are activated (Chen et al., 2005). The ERFs then bind to the cis-acting elements known as the GCC box (AGCCGCC, the core sequence of ERE) in the promoters of ethylene responsive genes, and activate these genes. DNA binding is mediated by the ERF AP2 domain (Ohme-Takagi and Shinshi, 1995; Ogawa et al., 2005). Recent data suggested that melon (*Cucumis melo*) ERFs were also induced by cold or drought stress conditions (Mizuno, 2006).

It has also long been recognized that plant hormone abscisic acid (ABA) plays an important role in adaptation to abiotic environmental stresses (Finkelstein et al., 2002).
The genes bearing ACGT core cis-acting elements in their promoter regions are regulated by ABA. The ACGT sequences responsive to ABA are referred to as ABA responsive elements (ABREs) (Busk and Pages, 1998; Rock, 2000). A conserved family of bZIP protein factors bind to the ABREs and are referred to as ABFs (ABRE binding factors) or AREBs (ABA-responsive element binding proteins) to activate gene expression (Choi et al., 2000; Uno et al., 2000). ABF1, ABF4, and maybe ABF2 are induced by cold treatment (Choi et al., 2000). ABA also plays a role in the activation of CRT elements through an uncharacterized ABA-inducible pathway by the function of CBFs (Knight et al., 2004). And ABA induces the expression of CBF1-3 (Knight et al., 2004). Thus the ABA-inducible pathway indeed interacts with ABA-independent pathway, the CBF (C-repeat binding factor) pathway.

Arabidopsis CBFs mediate plant cold tolerance and drought adaptation. The CBFs are also AP2 domain containing transcriptional activators. The AP2 domains of CBF and ERF are highly identical to each other (Hao et al., 2002). The CBF1 binding site contains the five nucleotide CCGAC (C-repeat, CRT; also called dehydration responsive element, DRE) core sequences (Baker et al., 1994; Yamaguchi-Shinozaki and Shinozaki, 1994). Using reporter construct harboring two CRT/DREs from the promoter of melon (Cucumis melo) 1-Aminocyclopropane-1-carboxylate (ACC) synthase (ACS, S-adenosyl-L-methionine methylthioadenosine-lyase) gene, recent investigation identified two CMe-ERFs and one CMe-DREB1 by the yeast one-hybrid strategy (Mizuno et al., 2006). This identification of ERFs and DREB suggests the identity of their AP2 binding domain.
Furthermore, *CMe-DREB1* and *CMe-ERFs* are activated under stress conditions (Mizuno et al., 2006).

The promoter regions of many stress inducible genes contain multiple *cis*-acting elements, suggesting the possibility of regulation by more than one signal transduction pathways. For example, the promoter region of *COR* genes contains more than one type of *cis*-acting elements. The *COR15a* contains three CRT/DREs and three ABREs (Baker et al., 1994). Thus transcription factors CBFs and ABFs can bind to their responsive *cis*-acting elements to activate downstream gene expression. Multiple *cis*-acting elements allow for synergistic activation (Giniger and Ptashne, 1988). More recent studies indicate that multiple binding sites and the potency of an activation domain augment the binding of an activator to its target promoter region (Tanaka, 1996). The *COR* genes in some crop species have more CRT/DREs when compared to Arabidopsis *COR* genes. For example, the promoter region of cold-inducible wheat *wcs120* contains at least eight CRT/DREs and five ACGT core elements (Vazquez-Tello et al., 1998). Our preliminary data indicated that ABF1 was induced by low temperature. *CBF1* is rapidly induced by low temperature, and then declines quickly, while *ABF1* is induced later. However, the *COR* genes continue to increase while the *CBFs* decrease. Thus, these two transcriptional activators with different activation domains contribute to the activation of *COR* genes. Different activators have the ability to transactivate gene synergistically, presumably by simultaneously interacting with the parts of transcriptional machinery (Lin et al., 1988; Ptashne and Gann, 1997; Sawada et al., 1999). For example, seed storage protein bZIP10 has a synergistic effect with ABI3 in a transient expression analysis (Lara
et al., 2003). Similar to bZIP10, the ABFs also belong to bZIP transcription factors. So, we hypothesized that the CBF and ABF with different characteristic nature had synergistic activity in the regulation of COR gene expression.

In this chapter, we investigated the connection between several different signal transduction pathways. Here, we present that CBF1 activates lacZ reporter gene driven by the natural COR15a promoter. Our constitutive phosphorylation mutant of ABF1 S104D failed to trans-activate reporter gene in yeasts suggests that additional modification is required for ABF1 to be fully functional.

4.3 MATERIALS AND METHODS

Yeast Reporter Strains

Yeast reporter strains EJS149, EJS216, and EJS222 were previously constructed by Stockinger (Stockinger et al., 1997). Reporter strain EJS222 contains two wild-type CRT/DREs in a tail-to-tail orientation; EJS149, two mutated CRT/DRE in a head-to-tail orientation; EJS216, two wild-type EREs in a tail-to-tail orientation (Figure 4.1 A).

To construct a reporter strain harboring a natural promoter, a 440 bp fragment of COR15a promoter region (Figure 4.2) was amplified from plasmid pEJS373, an EcoRI subclone of pSSB1 (Baker et al., 1994) using primers E100 (5′-GGAAGATCTTGTTCAACTCATTTGGGCTA-3′) and E101 (5′-GGAAGATCTGATCGATTACACTCCA-3′). PCR products were ligated to pGEM-T vector to create pZW65. Using incorporated Bgl II restriction enzyme sites (underlined nucleotides in primer
sequences), we cut and then ligated this fragment into the *Bgl* II site of the UAS-less “one-hybrid” vector pBgl-lacZ (Li and Herskowitz, 1993) to create pZW80. pZW80 and pBgl-lacZ vector were linearized with *Stu* I and then integrated into yeast “one-hybrid” host strain GGYI (Li and Herskowitz, 1993) creating yeast reporter strain ZW81 (Figure 4.4) and control ZW82 respectively. Assay used the antisense insertion of reporter constructs (Figure 4.4).

**ERF Activator Constructs**

The ERF genes do not contain introns (Arabidopsis Genome Initiative, 2000). To clone ERF1, ERF2 and ERF5, we amplified the respective genes from genomic DNA of Columbia ecotype using the following primers: ERF1 (ES52 5′- CGCGGATCCCGGATTCTCAATCTGA- 3′ and ES53 5′- CGCGGATCTCTTATA AAAACCAATAAACGATCGCCAC-3′); ERF2 (ES54 5′-CGCGGATCCATGTACGGA CAGTGCAATATAGAATC-3′ and ES55 5′- CGCGGATCCCTATGAAACCAATAACTCATCAACAC-3′); ERF5 (ES56 5′-CGCGGATCCCTAAC GAA GTATCTG-3′ and ES57 5′- CGCGGATCCTCAAACAAACGGTCAACTGGAATAAC-3′). PCR products of ERF1, ERF2, and ERF5 were ligated to pGEM-T easy vector as plasmid pZW1, pZW2, pZW3 respectively and confirmed by sequencing. We then used the incorporated *Bam* HI sites (underlined nucleotides in primer sequences) to subclone the *Bam* HI fragments of three ERFs into *Bgl* II site of yeast shuttle vector pDB20.1 (Berger et al., 1992), which was treated by Calf Intestinal Phosphatase (CIP, NewEngland
Biolab). The resulting plasmids were named pZW4, pZW5, and pZW6 respectively (Figure 4.1C).

**In Vitro Mass Excision of Phage**

A single colony of *E. coli* strain XL1-Blue MRF’ or *E. coli* strain SOLR was inoculated into 5ml of NZY supplemented with maltose (0.2% w/v). Cultures were grown at 37°C until OD$_{600}$ = 1.0, or at 30°C overnight. Cells were spun down and the pellets were then resuspended in 10 mM MgSO$_4$, using half volume of the culture medium. Resuspended cell pellets were diluted to OD$_{600}$ = 1.0. In a sterile culture tube, a portion of the amplified λ bacteriophage library CD4-15 or CD4-16 (ABRC Stock Center, Columbus, OH) and XL1-Blue MRF’ at an MOI (multiplicity of infection; i.e., ratio) of 1:10 λ phage particles per one cell were combined. ExAssist helper phage was added at an MOI of 10:1 (helper phage-to-cell ratio). The λ phage, helper phage, and *E. coli* XL1-Blue cells were incubated at 37°C for 15 minutes to allow complete infection by phage. Twenty ml LB was then added and incubated at 37°C for 2.5-3 h with vigorous shaking. The tube was then heated at 65-70°C for 20 minutes. The solution was spun down at ~ 1000g for 15 minutes. The supernatant, which contained the excised pBluescript phagemid packaged as filamentous phage particles, was transferred into a new sterile tube. The excised phagemid was titered using *E. coli* strain SOLR cells.

To prepare DNA from the *en masse* excised phagemids, a single colony of *E. coli* strain SOLR was inoculated into 10 ml LB until OD$_{600}$ = 1.0 at 37°C or 30°C overnight.
Then 5 ml of the excised phagemid (about 2~4 x 10^7 pfu) was added to the 10 ml culture of SOLR cells (roughly corresponding to 100x the original library obtained from ABRC). After incubation at 37°C for 15 minutes, these 15 ml cultures were added into 1000 ml LB supplemented with 150 µg/ml ampicillin. The cultures were grown until OD_600 = 1.0.

DNA was extracted using QIAGEN Plasmid Maxi Kit with 20 ml Buffer P1, P2, P3 respectively (QIAGEN, Valencia, CA). About 200 ng of DNA was used as a template to amplify the genes of interest.

**ABF Activator Constructs**

The ABF genes contain introns (Arabidopsis Genome Initiative, 2000). To obtain cDNAs for the four ABF genes, we designed primers based on database sequences to the entire coding sequence. The primers are as follows: ABF1 (ES98 5’-

CGCGGATCCATGGGTACTCACATTGATATC-3’ and ES99 5’-CGCGGATCCCTTACCACGGACCGGTAAGGG); ABF2 (ES100 5’-

CGCGGATCCATGGATGGTAGTATTTGG-3’ and ES101 5’-

CGCGGATCCCTCAACAGGTCCCGACTCTGTCC-3’); ABF3 (ES102 5’-

CGCGGATCCATGGGTCTAGATTAAACTTC-3’ and ES103 5’-

CGCGGATCCCTACCAGGGACCCGTCAATG-3’); ABF4 (ES104 5’-

CGCGGATCCGTCTTTGTTTTTCATTGAAAAAG-3’ and ES105 5’-

CGCGGATCCCTACCATGGTCCGTTAATGTCC-3’). *BamH I* sites were incorporated in these primers (underlined sequences in primers). The DNA prepared from the *en mass* excised cDNA libraries was used as template. After confirmation by
sequencing, these ABF coding sequences were ligated to pGEM-Teasy vectors, resulting in pZW34, pZW35, pZW36, and pZW37 respectively.

To generate yeast expression constructs, we first used Not I to digest pPC98, a yeast shuttle vector which harbored the yeast ADA2 gene, to remove the yeast ADA2 gene. The plasmid was allowed to self-ligate, yielding the original shuttle vector, and was named pZW63 (Figure 4.3). The ABF1-4 Bam HI fragments were then subcloned into the Bgl II site of pZW63, creating pZW66, pZW67, pZW68, and pZW69 respectively. Thus, each ABF gene is controlled by the promotor and terminator of acetoacetate decarboxylase gene (P ADC and T ADC). We then used Xho I and BamHI to remove the fragment of P ADC-ABF1-4 gene -T ADC, which was then ligated to the yeast low-copy-number shuttle vector pRS413 creating pZW70, pZW71, pZW72, and pZW73, respectively; or ligated to yeast high-copy-number shuttle vector pRS423 resulting in pZW74, pZW75, pZW76, and pZW77 respectively.

Frozen Yeast Competent Cell Preparation and Transformation

Frozen yeast competent cells were prepared using a protocol modified from the Gietz Lab (http://www.umanitoba.ca/faculties/medicine/biochem/gietz/Frozen.html). A single yeast colony was picked and grown in YPAD until the OD$_{600}$ reached 0.6 to 1.0. Cells were spun down at ~ 6000 g for 8 minutes. Harvested cell pellets were washed in Solution I {0.5 vol of 1.0 M. sorbitol, 10 mM Bicine-NaOH (pH 8.35), 3% ethylene glycol, 5% DMSO}. The cells were spun again and pellets were resuspended in 0.02X volume of the starting culture volume with Solution I. About 0.1 to 0.2 ml aliquots were
frozen and then stored at –80 °C until needed. It was noted (by the Gietz lab) that slow freezing results in good viability. Therefore, we covered the tubes with several layers of paper towel before putting into freezer.

For transformation, 0.5 - 5 µg of plasmid DNA and 50 µg of single stranded carrier DNA (20 mg/ml) in a maximum volume of 20 µl were added to the frozen yeast competent cells. Heat denatured Herring sperm DNA used for DNA and RNA gel blots can be used for this transformation. Yeast cells were then placed in a 37 °C water bath by tipping every 10 - 15 sec until the solution melted completely. Next, 1.4 ml of Solution II {40% PEG1000, 0.2 M Bicine-NaOH (pH 8.35)} was added and mixed by gently vortexing for about 1 min. Transformed cells were incubated at 30°C for 1 h. Yeast cells were spun down at 3000 x g for 5 sec and resuspended in 1.0 ml of Solution III {0.15 M NaCl, 10 mM Bicine-NaOH (pH 8.35)}. About 100 –200 µl resuspended cells were plated onto appropriate SC dropout plates.

Site Directed Mutagenesis and Resulting Plasmid

Primers used in PCR reaction to mutate ABF1 serine to aspartic acid (S104D) were as follows: S112 (coding sequence) 5’-GCCTAGAACGCTTgaTCAGAAGACTGTC-3’ and S113 (non-coding sequence) 5’-GACAGTCTTTCTGAtcAAGCGTAGATCC-3’. About 100 ng diluted pZW34 (pGEM-Teasy-ABF1) was used as template. The PCR conditions were as follows: 1 cycle at 95 °C for 5 min; then 18 cycles at 95 °C 1 min, 50 °C 1 min, and 68 °C 5 min; followed 1 cycle at 68 °C 10 min. Pfu Turbo DNA polymerase was used (Stratagene, La Jolla, CA). PCR products were first digested with
Dpn I. Dpn I only digests the parental plasmid DNA that has been N6-methylated the adenine residue and leaves the in vitro synthesized PCR products intact. PCR products were then transformed into E. coli DH5α competent cells. DNA from transformed colonies was extracted and the mutation was confirmed by sequencing. The resulting plasmid was named pZW240. BamH I fragment of ABF1_{S104D} from pZW240 was subcloned into Bgl II site of yeast vector pDB20.1 (Berger et al., 1992) as pZW233; into BamH I site of pCIB710 (Rothstein et al., 1987) as pZW229. Then Kpn I/Hind III fragment of pZW229 was ligated into binary vector pCIT20 (Ma et al., 1992) as pZW230.

4.4 RESULTS

β–galactosidase Activities of CBF and ERFs in Different Yeast Reporter Stains Harboring CRT/DREs or EREs.

To determine whether CBFs (C-repeat binding factor) might bind to EREs and similarly to determine if ERFs (ethylene-responsive element binding factor) bind the CRTs, we isolated three ERFs and cloned them into a yeast expression vector (Figure 4.1C). The expression of CBF1 in yeast reporter strain EJS222 confirmed that CBF1 activated the expression of CRT/DRE-responsive-lacZ reporter genes (Figure 4.1 A and B). In contrast, β–galactosidase assays indicated that ERF1, ERF2, and ERF5 (pZW4-6) did not activate the CRT/DRE-responsive-lacZ reporter gene (Figure 4.1 B). However, ERF1, ERF2, and ERF5 also did not activate the expression of the ERE-responsive-lacZ
reporter gene and had a similar activity level as the vector control (Figure 4.1 A and B). The CBF1\textsubscript{1-115} fragment (DNA binding domain only) had a background level of β−galactosidase enzymatic activity. Similarly, the CBF1\textsubscript{1-115}/VP16\textsubscript{413-490} fusion did not result in activation of reporter gene lacZ (Figure 4.1 B).

None of the activator constructs activated the reporter gene under control of the mutated CRT/DRE elements (Figure 4.1 B).

Yeast reporter strain EJS216 contains the two wild-type EREs in a tail-to-tail orientation (Figure 4.1A). Expressions of the different ERF constructs resulted in β−galactosidase enzymatic activity. Meanwhile, expression of CBF1\textsubscript{1-115}, and CBF1\textsubscript{1-115}/VP16 also resulted in β−galactosidase activity. However the control vector pDB20.1 also resulted in similar β−galactosidase activity, suggesting a leaky expression. On the contrary, CBF1 expression might have a negative effect on the leaky expression of lacZ gene (Figure 4.1). From above experiments, we conclude that CBF1 activates lacZ reporter genes bearing two copies of CRT/DREs in the promoter. Furthermore, lacZ gene has a leaky expression bearing two EREs.

β−galactosidase Activities of CBF and ABFs in Yeast Reporter Stain Harboring the Natural COR15a Promoter

To determine whether CRT/DRE and ABRE binding factors might act together upon the COR gene promoter to give greater expression than either factor alone, we created a native COR promoter lacZ fusion ZW81 (Figure 4.4 A). The transformation of
the ABF1-4 constructs and the CBF1 activator construct into yeast reporter strain ZW81 indicated that CBF1 activated lacZ reporter gene, whereas none of the ABF1-4 constructs activated gene expression (Figure 4.4 B). Co-transformation of CBF1 with the four different ABFs did not increase the $\beta$–galactosidase activity (Figure 4.4 B).

**Constitutive Phosphorylation Mutant ABF1$_{S104D}$**

Recent results indicate that Arabidopsis ABF homologs in rice TRAB1 (transcription factor responsible for ABA regulation) and wheat TaABF require phosphorylation to function (Johnson et al., 2002; Kagaya et al., 2002). Thus, we reasoned that the Arabidopsis ABFs might also need similar post-translational modification to trans-activate. The critical residue of rice TRAB1 is Ser102 (Kagaya et al., 2002). Alignment of the AtABF1-4 with TRAB1 indicated that Ser104 of ABF1 was probably the conserved residue (Figure 4.5). Primers were designed to change Ser104 of ABF1 to Asp. This mutation is expected to produce a constitutive phosphorylation mimic. The ability of ABF1$_{S104D}$ to transactivate was tested in yeast reporter strain ZW81. This mutant also failed to activate the COR15a-lacZ reporter gene. Thus we conclude that the S104D mutation of ABF1 is still insufficient for ABF1 to activate in yeast.

**4.5 DISCUSSION**

The expression of CBF1 in yeast reporter strain EJS222 with two copies of CRT/DRE activates lacZ reporter gene, whereas the expression of CBF1$_{110-115}$/VP16$_{413-490}$
did not. We knew that VP16\textsubscript{AD} could be a functional AD in yeast as indicated by our GAL4/VP16\textsubscript{AD} fusion. One difference is that GAL4 fusion can form dimers, whereas CBF1\textsubscript{DBD}, its AP2 domain, possibly act as a monomer (Hao et al., 2002). This difference might imply that there is a threshold for VP16\textsubscript{AD} to function. These experiments suggest that there are different threshold requirements for a given transcriptional activator; it might be dependent on the potency of activation domain. Observation tells us that usually one copy of a cis-acting element is not good enough for constructing a yeast reporter strain. Our data suggest it is safe to add more than two copies of cis-acting elements in reporter construct when making yeast one-hybrid reporter strains.

We were surprised to find out that our expressed ABFs in yeast did not activate the downstream reporter gene expression, because ABFs were originally identified by using the yeast one-hybrid strategy (Choi et al., 2000; Uno et al., 2000). Further checking the transactivation assays for ABFs done by Choi et al. indicated that the trans-activity resulted from ABFs was not significantly high than the control (Choi et al., 2000). In-gel protein kinase assay indicated that the conserved domain (putative activation domain) of ABFs were phosphorylated in a ABA-dependent way (Uno et al., 2000). Thus we reasoned that ABFs might need a posttranslational modification in order to be fully functional in yeast, as (de-)phosphorylation is a frequent mechanism used by cells to regulate transcription factor activity (Whitmarsh and Davis, 2000). This may be particularly true for bZIP factors. Recent publications reinforced this hypothesis, as indicated by rice homologous bZIP factor, TRAB1, and a parsley bZIP factor. Both factors’ binding activities are dependent on phosphorylation (Wellmer et al., 2001;
Kagaya et al., 2002; Kobayashi et al., 2005). Another line of evidence from barley research that an ABA-responsive kinase PKABA1 phosphorylates TaABF by interacting with it, strengthens this hypothesis (Gomez-Cadenas et al., 1999; Johnson et al., 2002). However, our S104D mutation of ABF1 did not activate reporter gene in yeast suggested that ABF1 might need additional modification in order to be functional. Results with analysis of ABF2 supported this idea that ABA-induced further modification was needed to activate downstream COR genes in transgenic ABF2 lines (Fujita et al., 2005).

4.6 REFERENCES


AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. Plant Cell 17, 3470-3488.


Figure 4.1 β-galactosidase levels of transcriptional activators in reporter strains harboring different cis-acting elements.

(A) Three yeast reporter strains used in the experiments. EJS222 harbors two CRT/DREs in a different orientation; EJS149, two mutated CRT/DREs in a same orientation; EJS216, two EREs in a same orientation.

(B) β-galactosidase levels of different constructs in three reporter strains EJS222, EJS149, or EJS216 respectively. Five independent colonies were cultured and assayed for enzymatic activity. The errors are standard deviation of mean.

(C) Schematic represent of different effector constructs used in β−galactosidase assay. ERF1, 2, and 5 (pZW4, 5, and 6), CBF1 full length (pEJS251), CBF1<sup>1-115</sup> (pEJS447), and CBF1<sup>1-115</sup>VP16<sub>413-490</sub> (pEJS448) are shown. All these genes were put under the control of the yeast alcohol dehydrogenase (ADH) gene promoter and terminator.
Figure 4.2 Sequences upstream of TATA box in COR15a promoter region that harbors three CRT/DREs and three ABREs were used to construct yeast reporter strain ZW81.

CRT/DREs (CCGAC) are highlighted in blue and underlined; ABREs, ABA-responsive element, are in Italic and underlined; TATA box, in bolded red and underlined. Last three residues are translation initiation start site ATG (pink). The transcription start site A (highlighted in green and bolded) is as +1. Three CRT/DREs (CCGAC or CAGCC) are located at positions around -420, -357, and -180 respectively, whereas the three ABREs with the core sequence ACGT locate around -305, -127, and -68 respectively. Sequences just prior to TATA box were used to construct yeast reporter strain.
Yeast ADA2 gene was removed from pPC98, and self-ligated to recover the original yeast shuttle vector, named as pZW63. ABF1-4 genes were ligated into Bgl II site of pZW63. Then Xho I and BamH I fragment containing P ADC, ABF, and T ADC (P ADC and T ADC, promoter and terminator of acetoacetate decarboxylase gene) was subcloned to responsive site of yeast shuttle vector pRS413 (low copy expression vector) or pRS423 (high copy expression vector).
Figure 4.4 β-galactosidase level of ABF and CBF transcriptional activators in yeast reporter strain ZW81 harboring COR15a promoter.

(A) A construct harboring three CRT/DREs (arrows) and three ABREs (rectangle) was integrated into yeast genome creating reporter strain ZW81.

(B) β-galactosidase levels of transcriptional activators in ZW81. ABF and CBF coding regions were under the control by alcohol dehydrogenase gene promoter and terminator. ABF L = single copy yeast plasmid (cen). ABF H = high copy yeast plasmid (2 μ). For coexpression, CBF1 and ABFs either in pRS413 (ABFsL) or pRS423 (ABFsH) were cotransformed into ZW81.
<table>
<thead>
<tr>
<th>TRAB1</th>
<th>ABF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAB1 1</td>
<td>ABF1 1</td>
</tr>
<tr>
<td>TRAB1 11</td>
<td>ABF1 61</td>
</tr>
<tr>
<td>TRAB1 62</td>
<td>ABF1 117</td>
</tr>
<tr>
<td>TRAB1 116</td>
<td>ABF1 173</td>
</tr>
<tr>
<td>TRAB1 209</td>
<td>ABF1 293</td>
</tr>
<tr>
<td>TRAB1 269</td>
<td>ABF1 351</td>
</tr>
</tbody>
</table>

Figure 4.5 Alignment of Arabidopsis ABF1 and rice TRAB1.

The Ser residue requiring phosphorylation in TRAB1 (S102 in Kagaya et al.) is in bold, italic, and underlined. The corresponding Ser 104 residue in Arabidopsis ABF1 is also in bold, italic, and underlined.
CHAPTER 5

METHYLATION MUTANTS HARBORING MULTIPLE COPIES OF A TRANSGENE WITHIN ONE LOCUS EPIGENETICALLY REGULATE THE EXPRESSION OF GENES INVOLVED IN PLANT COLD ACCLIMATION

5.1 ABSTRACT

Epigenetic control of transcription plays a pivotal role in tissue-specific gene expression and gene silencing. We determined the role of epigenetic control of transcription on the expression of genes related to plant cold acclimation. Several methylation mutant lines were subjected to low temperature in a time course treatment. RNA gel blot analyses revealed that the integration of multiple copies of HPT gene into one chromosomal locus dramatically reduced the expression of CBF1, COR15a, and ABF1, suggesting a link between the integrated repeats and transcriptional regulation of genes involved in plant cold acclimation. Plant freezing assays revealed that line som 8 and mom + were indeed more sensitive to freezing temperature than line mom – and wild type. Thus, our results reveal that endogenous genes involved in plant cold acclimation are also affected by integrated multiple transgene repeats. The transcript levels of HPT
gene in som 8 were several folds higher than mom+ lines, suggesting significant effect of demethylation to release TGS (transcriptional gene silencing) compared to the role played by MOM without any demethylation. This observation also suggests that the silence nature of transgene locus is not required to trigger epigenetic regulation of endogenous gene expression; that SOM and MOM might act in a different pathway. Based on the current understanding of the effect on TGS by duplicated repeats, I propose a model that the chromatin structure created by duplication and repeats in the integrated transgene locus triggers an unknown silencing mechanism, which achieves the epigenetic regulation of genes related to low temperature tolerance. This research provides new insights into how multiple copies of a transgene has an epigenetic regulation of endogenous gene expression.

5.2 INTRODUCTION

Epigenetic modifications are heritable but reversible alterations in gene expression that occur without the change in the primary DNA sequence. Epigenetic modification plays a pivotal role for plants and animals to achieve tissue-specific expression of exogenous DNA sequences, such as transgene, transposons, and viruses (Fagard and Vaucheret, 2000; Chan et al., 2005; Grant-Downton, 2005). Gene silencing may occur at the post-transcriptional or transcriptional level. Transcriptional gene silencing (TGS) leads to the loss of gene expression by blocking transcription initiation, whereas post-transcriptional gene silencing (PTGS) results from the degradation of transcribed RNA.
(Vaucheret and Fagard, 2001). TGS can occur either in cis or in trans (Fagard and Vaucheret, 2000). TGS in cis results from the adjacent silencing structure. For example, an inhibitive effect is created by the neighboring silenced transgene repeats on neighboring heterochromatin (Fagard and Vaucheret, 2000). TGS can also occur in trans, either through a transient DNA-DNA pairing between the unlinked locus and the target genes, which result in the trans-methylation of the target gene by the structure(s) at the heterochromatin, or a specific molecule resulted from silenced loci (Fagard and Vaucheret, 2000).

To understand the molecular mechanism of TGS, plant biologists utilized silenced transgene loci. The locus in question contains multiple copies of a marker gene, (Mittelsten Scheid et al., 1991; Matzke et al., 1994; Kumpatla and Hall, 1998). In one case, a CaMV35S promoter HPT (hygromycin phosphotransferase) gene construct was introduced into Arabidopsis (Mittelsten Scheid et al., 1991; Mittelsten Scheid et al., 1994). T1 transgenic plants were all hygromycin resistant. However, the T2 plants from one selfed line (line A) were all hygromycin-sensitive. Further analyses revealed that multiple copies of the HPT gene had integrated into a single locus, and that these silenced transgenes were hypermethylated (Mittelsten Scheid et al., 1996). Subsequent mutagenesis of line A resulted in eight mutants som1-8 (for the somniferous effect of the wild type allele) that restored the hygromycin resistance (Mittelsten Scheid et al., 1998). In these mutants, the HPT transgenes were no longer hypermethylated. Allelism tests and sequencing suggested that som1, 4, 5, 6, 7, and 8 might be allelic to ddm1 (decrease in DNA methylation) (Vongs et al., 1993; Mittelsten Scheid et al., 1998; Jeddeloh et al., 1998).
DDM1 encodes a protein related to SWI2/SNF2 chromatin-remodeling protein factors (Jeddeloh et al., 1999). DDM1 is necessary to maintain the TGS in Arabidopsis (Jeddeloh et al., 1998). Disruption of DDM1 reduces the cytosine methylation of the Arabidopsis genome by 70%, primarily in the repeated sequences (Vongs et al., 1993). Low-copy sequences are also affected but the methylation is gradually lost after selfing for multiple generations (Jeddeloh et al., 1998).

Methylation plays an important role in gene regulation (Bird and Wolffe, 1999). Although methylation is often associated with the gene silencing, recent investigations suggest that the methylation of a gene does not necessarily imply silencing. Similarly, demethylation of a gene does not necessarily imply activation (Tweedie and Bird, 2000).

Hygromycin sensitive line A was also mutated by random T-DNA insertion. One hygromycin resistant line was identified but the hypermethylation pattern of line A was maintained. Map based cloning of the responsible gene led to the identification of MOM (for ‘Morpheus molecule’) gene, which encodes a protein predicted to have ATPase activity and chromatin remodeling activity (Amedeo et al., 2000). MOM is most similar to that of the SWI2/SNF2 family (Amedeo et al., 2000). Arabidopsis plants defective for MOM1 release TGS (Chandler and Jorgensen, 2000; Tariq et al., 2002). The mom1 mutants reactivate HPT expression but maintain dense methylation of the HPT gene cluster at levels identical to the progenitor line (line A) even after nine generations of selfing (Amedeo et al., 2000). These data suggest that the methylation status is not necessarily correlated with transcriptional silencing activity. The reactivated HPT gene cluster is thought to result from a chromatin structural change, rather than demethylation.
It has been established that gene expression changes in plants upon low temperature treatment. Plants can increase in freezing tolerance after a period of exposure to low temperature, a process known as plant cold acclimation (Thomashow, 1999). Cold acclimation involves gene expression changes, which can be separated into two distinct pathways, including an ABA independent pathway, and an ABA dependent pathway (Thomashow, 1999; Xiong et al., 2002; Yamaguchi-Shinozaki and Shinozaki, 2005).

The ABA independent pathway is mediated by a small family of transcriptional activator CBFs, C-repeat binding factors (Stockinger et al., 1997; Jaglo-Ottosen et al., 1998). There are three CBFs, CBF1, CBF2, and CBF3, that have similar roles in the activation of plant cold-regulated (COR) genes involved in plant cold acclimation (Gilmour et al., 2004). Most COR genes contain two different cis-acting regulatory elements, the C-repeat/dehydration responsive element (CRT/DRE, CCGAC), and the ABA responsive element (ABRE, PyACGTGGC). The Arabidopsis COR15a gene promoter region harbors three CRT/DREs and three ABREs (Baker et al., 1994). The CRT/DREs confer ABA independent regulation of gene expression, whereas ABREs confer ABA dependent regulation of gene expression (Choi et al., 2000). Low temperature stimuli results in elevated endogenous ABA levels in plants, which leads to the induction of four ABF (ABRE binding factor) transcription factors (Choi et al., 2000; Uno et al., 2000). ABF1 expression is highly induced by cold treatment, while ABF4 expression is cold induced but to relative low levels (Choi et al., 2000). All four ABFs are induced by ABA treatments (Choi et al., 2000). All four ABFs also bind to the ABREs present in COR.
gene promoters, and can activate COR gene expression. These ABFs mediated the ABA dependent pathway (Choi et al., 2000; Kang et al., 2002).

Interestingly, low temperature also plays a role in demethylation. Burn et al. (1993) proposed that prolonged low temperature treatment (vernalization) reduced the methylation level of certain Arabidopsis genes involved in flowering time determination, which then permitted their transcription. Indeed, the decrease of DNA methylation level promoted the flowering in Arabidopsis even without the need of cold treatment (Finnegan et al., 1998). DNA demethylation has also been observed to occur after cold treatment. In these experiments, the tissue type and plant age of the treated plants (Cichorium intybus L.) are important factors (Demeulemeester, 1999). In maize, cold treatment selectively demethylates the DNA in the region of the Ac/Ds transposon locus, and primarily in the maize roots but not the shoots (Steward et al., 2000). Thus, epigenetic modification is a strategy that plants use to achieve tissue-specific gene expression.

However, whether epigenetic modification plays a role in the regulation of genes involved in plant cold acclimation remains unknown. Here, we hypothesized that epigenetic control of transcription might play a role in the expression of genes important for cold acclimation. To test this hypothesis, we utilized Arabidopsis methylation mutants, conducted time course cold treatments and determined the freezing tolerance of the mutants. Our findings indicate that the integration of multiple copies of a transgene into a single locus negatively affects the expression of CBF1, ABF1, and COR genes. This negative effect might be associated with the organization of the transgene locus but not the silence nature of the transgene. Thus, our investigation reveals that plants adopt
an epigenetic strategy to regulate endogenous genes, which are involved in plant stress tolerance.

5.3 MATERIALS AND METHODS

Plant Materials

Arabidopsis lines used in this research were obtained from O. Mittelsten Scheid (Friedrich Miescher Institute, Switzerland). A description of how Scheid and colleagues created these lines follows. Arabidopsis ecotype Zürich was transformed with the hygromycin phosphotransferase gene (HPT) under the control of the CaMV35S promoter (Mittelsten Scheid et al., 1991). Hygromycin resistant transgenic seedlings were selfed and brought to homozygosity. One line, line A, harbored multiple copies (5 to 10) of the 35S:HPT transgene at a single locus but was hygromycin sensitive. Subsequent analyses indicated that the HPT transgenes were silenced (Mittelsten Scheid et al., 1994). The line A was mutagenized by ethyl methane sulfonate (EMS) or by fast neutron radiation (FNR). The screening of the mutagenized M2 population and selection for hygromycin resistance identified eight independent mutations, som1-8. The methylation levels of HPT genes, as well as at repetitive sequences in som 8 plants were found to be similar levels of those in ddm1 mutant (Mittelsten Scheid et al., 1998). Gene som 8 is allelic to ddm 1 (Jeddeloh et al., 1999).

The mom1 mutants were identified using random T-DNA insertional mutagenesis in line A, and screening for hygromycin resistance and hypermethylation pattern of line A.

128
(Amedeo et al., 2000). Thus, mom + is hygromycin resistant and the HPT repeats are hypermethylated. The mom - line was created by backcrossing to the wild type Zürich, and crossing out the HPT locus but maintaining the mom T-DNA insertion.

**Low Temperature Treatment**

Seedlings from the different lines were grown under 16h light/8h dark at 22 °C for approximately 4 weeks. Plants were then placed at 4 °C and harvested at different times for the extraction of total RNA. Low temperature treated seedlings were separated into two pools for 0h, 4h samples and one pool for 6h samples.

**Plant Freezing Assays**

Arabidopsis seeds were surface sterilized and cold stratified at 4 °C for four days. About 20 seeds were sown onto each plate filled with Gamborg B5 media. A total of 5 plates per line were tested. Germinated seedlings were grown on Gamborg B5 plates at 22 °C with 16 h light/ 8 h dark cycle for two weeks. Plates were then transferred to –2 °C and constant darkness for 3 h. After which, the covers were removed and the plants were misted with sterile cold water to initiate ice nucleation. Plates were kept at –2 °C for additional 21 h before decreasing the temperature to –6 °C for 24 h. Plates were then moved to +2 °C and maintained in the dark for 8 h. The temperature was then increased to 4 °C for 16 h before moving plants back to 22 °C with 16 h light/ 8h dark cycle for 20 h prior to photography.

**RNA Extraction and RNA Gel Blot Analyses**

Cold treated plants were harvested in pools and immediately frozen in liquid nitrogen. Tissues were then ground in liquid nitrogen. Total RNAs were extracted using
Plant RNeasy mini kits (Qiagen Inc., Valencia, CA). Ten µg of total RNA was fractioned on a 1% formaldehyde gel and blotted to Hybond N+ (Amersham Biosciences Corp., Piscataway, NJ). DNA probes were radiolabeled with $^{32}$P-nucleotides using Random-Prime labeling kits (Amersham Biosciences Corp., Piscataway, NJ). The HPT gene probe was generated by PCR amplification of 864 bp fragment of HPT gene from binary vector pCIT20 (Ma et al., 1992) using oligos EJS41 5´-
GTCCTGCGGGTAAATAGCTGCGCC-3´ and EJS42 5´-
GCTGGGGCGTCGGTTTCCACTACTGG-3´ (Primers HM+ and AP1 respectively in Mittelsten Scheid et al. (Mittelsten Scheid et al., 1994). The CBF1 gene probe was generated using primers EJS39 (M13 forward) 5´-
CGCCAGGGTTTTCCCAGTCACGAC-3´ and EJS40 (M13 reverse) 5´-
AGCGGATAACAATTTCACCAGGA-3’, which produced a full length CBF1 pEJS114 (Stockinger et al., 1997). Full length fragments of ABF1, eIF-4A genes were used for probe labeling. Filters were hybridized, stripped, and re-hybridized sequentially with probes for COR15a, CBF1, HPT, ABF1, and eIF-4A. RNA blot results were visualized by phosphorimage autoradiography using a Storm 860 PhosphorImager (Amersham Pharmacia Biotech Inc., Piscataway NJ). Hybridization signal bands of targeted genes and a constitutive gene eIF-4A were quantified. Message abundance of each gene was represented as the relative expression (target gene/eIF-4A) during a low temperature time course.
5.4 RESULTS

Low Temperature Induction of COR15a Gene Is Reduced in Line A, Som 8, and Mom+

To determine the role that different methylation mutants might play in the regulation of COR genes during plant cold acclimation, we subjected the methylation mutants line A, som 8, mom +, and mom - plants in the Zürich background to a low temperature time course cold treatment. RNA gel blot analyses revealed that COR15a gene expression was negatively affected in the methylation mutants line A, som 8, and mom +, when compared with wild type Zürich (Figure 5.1 A and C). The methylation mutants line A, som 8, and mom + also had a more pronounced effect in the reduction of COR15a gene expression at 6 h treatment than 4 h treatment (Figure 5.1 A and C). Comparison among line A, som 8, and mom + indicated that som 8 lines had the lowest COR15a gene message abundance at the 6 h time point (Figure 5.1 A and C). Thus we conclude that the regulation of COR15a is negatively affected in the three methylation mutants line A, som 8, and mom +.

Transcript Levels of CBF1 Are Reduced in Methylation Mutants Line A, Som 8, and Mom+

RNA gel blot indicated that COR15a expression was lower in line A, som 8 mutants, and mom+ mutants compared to the Zürich wild type parent (Figure 5.1). The CBF transcriptional activators regulate COR gene expression (Stockinger et al., 1997; Gilmour...
et al., 2004; Wang et al., 2005). Thus, we asked whether the transcript levels of CBF1 were similarly affected. RNA gel blot results indicated that CBF1 transcript levels in line A, som 8, and mom + were significantly reduced at both the 4 h and the 6 h time points (Figure 5.1 A and B). Line A had the lowest level of CBF1 expression compared to som 8 and mom + (Figure 5.1 A and B). The mom – mutant did not show any difference with Zürich in CBF1 expression. From these results, we conclude that CBF1 expression is also negatively affected in the methylation mutants, line A, som 8, and mom +.

**Line A and Som 8 Have Reduced Levels of ABF1**

RNA gel blot indicated that COR15a expression was negatively affected in the three methylation mutants, line A, som 8, and mom +. The promoter region of COR15a contains three CRT/DREs and three ABREs (Baker et al., 1994). The CRT/DRE is the binding site of the CBFs, the ABRE is the binding site of the ABFs. To determine whether ABF expression was similarly affected in the methylation mutant lines, we hybridized ABF1 to the filter (Choi et al., 2000). ABF1 expression level in line A and som 8 was reduced at both the 4 h and the 6 h time points (Figure 5.1 A and D). In contrast to CBF1 and COR15a, ABF1 expression was not negatively affected in mom + at any time point when compared to Zürich wild type. The mom- mutant had no significant effect on the regulation of ABF1 gene expression. Thus, we conclude that the line A, and som 8 have reduced ABF1 expression, but mom+ and mom- did not.
Transcript Levels of HPT Are Higher in mom + than in som 8

The HPT genes in lines som 8 and mom + are reactivated after mutagenesis (Mittelsten Scheid et al., 1998; Amedeo et al., 2000). To confirm this result in our lab, we hybridized the filters with HPT. As expected, HPT gene was only expressed in line som 8 and mom + (Figure 5.1 A and E). RNA gel blot results revealed that the transcript level in line A was higher than in som + (Figure 5.1 A and E). This analysis also suggests that the silence nature of the transgene locus is not necessary to trigger epigenetic regulation of genes involved in plant cold acclimation.

Methylation Mutants Are More Sensitive to Freezing

To determine whether the methylation mutants were affected in freezing tolerance, we conducted plant freezing assays first with nonacclimated seedlings. Plants were taken from 22 °C to –2 °C for 24 h, and then to –6 °C for 24 h. Twenty hours after returning plants to normal growth conditions, lines som 8 and mom + showed sensitivity to freezing temperature, with som 8 more severe than mom + (Figure 5.2). In contrast, the mom – and wild type Zürich were unaffected by this freezing treatment. Thus, we conclude that methylation mutant lines som 8 and mom + are less freezing tolerant than the wild type Zürich plants under the non-acclimating conditions.
5.5 DISCUSSION

Epigenetic modification plays a pivotal role for plants and animals to achieve tissue-specific gene expression and gene silencing (Amedeo et al., 2000; Chan et al., 2005). Epigenetics plays a role in the regulation of genes involved in flowering time determination and flower development (Chan et al., 2005). Our objective in this experiment was to test the hypothesis that the low temperature induction of genes in the CBF response pathway might be altered in the methylation mutants. To that end, we employed several mutant lines affecting transcriptional gene silencing. RNA gel blot analyses indicated that CBF1 transcript levels were dramatically reduced in these methylation mutants, line A, som 8, and mom + compared to wild type Zürich but not in mom - (Figure 5.1). Similarly, COR15a expression in these three lines was also reduced but not in mom -. ABF1 was also reduced line A and som 8. In line mom-, the HPT locus was crossed out. The reduced expression of CBF1, COR15a, and ABF1 in line A, som 8, and mom+ (ABF1 not affected in this line) but not in mom- suggests that multiple copies of HPT gene in one single locus have a negative effect on the expression of these cold inducible genes, which are located on different chromosomes (chromosome 4, 2, and 1 respectively) with major roles in plant cold acclimation.

Plant freezing assays revealed that som 8 and mom + were more sensitive to freezing temperature than wild type Zürich (Figure 5.2), with som 8 being the most sensitive to freezing temperature than line mom+. Interestingly, som 8 plants also had the lowest COR15a expression levels of all of the lines examined (Figure 5.1 A and C).
The *mom*– line exhibited a similar phenotype as with wild type, and showed no effect on freezing tolerance. Thus, our RNA gel blot analyses and plant freezing assays suggest that the negative affection on the expression of these cold inducible genes is associated with the transgene locus.

Because the *HPT* gene within the transgene locus is reactivated in lines *som8* and *mom +* (Figure 5.1 A and E), it indicates that the silence nature of the transgene locus is not required to trigger the negative regulation of these cold inducible genes. Many previous investigations use a silenced locus to address the silencing of the ectopic target transgenes (Matzke et al., 1994; Thierry and Vaucheret, 1996; Yang et al., 2005). However, the underlying mechanism is not clear. It is proposed that the mechanism may involve transient DNA-DNA pairing to transmit the silence from loci to their targets or involve the production of silencing molecule from the silencing loci (Fagard and Vaucheret, 2000; Muskens et al., 2000). Our results suggest that it might be the chromatin structure or organization resulted from the rearranged transgene but not the silencing nature to trigger the epigenetic regulation. Thus, our investigation provides a new case of epigenetic regulation.

Multiple repeats of an integrated transgene can transcriptionally silence target genes. How the multiple repeats transmit TGS to its target genes remains unknown. It was suggested that this TGS could occur either in *cis* or in *trans* (Fagard and Vaucheret, 2000; Vaucheret and Fagard, 2001). TGS in *cis* may result from an adjacent silencing structure. For example, the inhibitive effect might be created by the neighboring silenced transgene repeats or the neighboring heterochromatin (Fagard and Vaucheret, 2000). TGS could
also occur in \textit{trans}, either a transient DNA-DNA pairing between the unlinked loci and the target genes resulting in \textit{trans}-methylation mechanism or a specific molecule resulted from silenced loci imposing inhibitive effect on target genes (Fagard and Vaucheret, 2000). The location of the multiple \textit{HPT} gene copies in line A and its derivative lines is not known. Therefore, we cannot rule out the possibility that \textit{HPT} locus acts in \textit{cis} by creating a silencing structure that blocks neighbor gene expression. However, affected \textit{CBF1} and \textit{ABF1} genes are located on chromosome 4 and 1 respectively, so a \textit{cis}-acting model seems unlikely. We think the reduction of \textit{COR15a} results from the reduced \textit{CBF1} and \textit{ABF1} but not from TGS. Thus TGS in \textit{cis} might not be our case. We therefore favor the \textit{trans} model in which the silenced locus transmits TGS to target genes (Fagard and Vaucheret, 2000). As described previously, we think that chromatin structure or organization resulted from the transgene locus triggers an unknown mechanism to epigenetically regulate (in \textit{trans}) these cold inducible genes, which are located on different chromosomes.

Our results suggest that the multiple repeats in several methylation mutant lines repress genes involved in plant cold acclimation in \textit{trans}. But how might multiple repeats of a transgene in one locus repress cold induced gene expression? First, we focused on the transgene locus. In line A, the silenced \textit{HPT} genes and genomic repetitive sequence were hypermethylated. In \textit{som 8}, the reactivated \textit{HPT} genes and genomic repetitive sequences were no longer hypermethylated and reduced to the levels of \textit{ddm1} mutants, whereas \textit{mom+} \textit{HPT} genes and repetitive sequences were still methylated but \textit{HPT} genes were reactivated. It is not known the changes of methylation levels of these
cold inducible genes. But, these analyses reveal that the silence of \textit{HPT} gene is not associated with their methylation status and might be resulted from the altered chromatin structure in the repeated region to cause the silence. So we reasoned this might also be the truth for epigenetic regulation of these cold inducible genes’ expression as these three lines had similar results (Figure 5.1). Line A contains multiple copies of the \textit{HPT} gene, which was rearranged before integration into one locus (Mittelsten Scheid et al., 1994). Presumably this rearrangement resulted in inverted repeats and/or duplicated sequences, which may act similar to transposon. And these inverted repeats and duplicated segments could trigger silencing (Henikoff, 1998). Additionally, at least one endogenous gene family, encoding Tryptophan enzyme phosphoribosylanthranilate isomerase (PAI), was shown to be epigenetically regulated (Bender and Fink, 1995). A tail-to-tail inverted duplication of PAI in one locus triggered the cytosine methylation of two additional unlinked PAI genes over the regions of DNA identity (Luff et al., 1999; Melquist et al., 1999). The \textit{CBF} genes also show a direct repeat pattern of “Promoter-CBF-Promoter-CBF-Promoter-CBF” with the highest identity in \textit{CBF} genes (Figure 5.3). Three \textit{CBF} genes are intronless and are 87% identical in the nucleotide coding sequences. But it is obvious \textit{CBF} genes are different from \textit{HPT} gene in the coding and promoter regions. So it is not similar to the case of PAI. Furthermore, the \textit{HPT} locus also negatively affects the expression of \textit{ABF1} expression, and \textit{ABF1} donot have the repeat pattern (Figure 5.1 A and C). Thus, it might be not the results from the methylation of homologous sequences or from the repeat pattern (Matzke et al., 1994; Jakowitsch et al., 1999). Based on these observations, I proposed a model that the specific chromatin structure resulting from the
transgene locus rather than its silencing status triggered unknown epigenetic mechanism to genome-widely regulate the expression of genes involved in cold tolerance (Figure 5.2).

Higher transcript levels of \textit{HPT} in \textit{som 8} than in \textit{mom+} suggest that the demethylation may have a more profound effect on the relieving TGS than the function of MOM. Tweedie and Bird proposed two silencing mechanisms for the MOM protein (Tweedie and Bird, 2000). The first scenario is that MOM mediates TGS directly downstream of DDM1 methylation, a function similar to mammalian methyl-CpG-binding protein. The methyl-CpG-binding protein binds to methylated DNA and recruits associated silencing proteins, which alter chromatin structure through the deacetylation of histones (Bird and Wolffe, 1999; Tweedie and Bird, 2000). The second scenario is that MOM and DDM act independently. Our Northern results favor the second scenario as suggested by the discovery of higher \textit{HPT} transcript levels in \textit{som 8}, in comparison to \textit{mom+} mutant line.

Intriguingly, low temperature also plays a role in PTGS. The inhibition of PTGS pathway by low temperatures is observed in insects, plants, and mammalian cells (Fortier and Belote, 2000; Szittya et al., 2003; Kameda et al., 2004). The PTGS pathway is inhibited by the control of small interfering (si) RNAs generation, which may explain the reason for the outbreaks of plant virus diseases under low temperatures. Under 15 °C, the generation of siRNA is undetectable (Szittya et al., 2003). Previous research has established that low temperature treatment can demethylate or methylate genes depending on the plant age and tissues (Demeulemeester, 1999; Steward et al., 2000). In this
research, our analyses suggest that the chromatin structure or organization created by a transgene locus can epigenetically regulate genes induced by low temperature. Thus, our results extend the list regulated by epigenetics beside tissue-specific gene regulation or genes involved in development. We summarized these new findings in table 5.1. Furthermore, our analyses and previous investigations in TGS and PTGS under low temperature as described above suggest a particular role played by low temperature, as one of the important environmental factors in plant life cycle, for plants to achieve epigenetic regulation of gene expression.

5.6 ACKNOWLEDGMENTS

We thank Dr Ortrun Mittelsten Scheid (Friedrich Miescher Institute, Switzerland) for providing the Arabidopsis lines used in this research, and Dr. Joyce Pennycooke for help with freezing assays.

5.7 REFERENCES


Steward, N., Kusano, T., and Sano, H. (2000). Expression of ZmMET1, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in actively proliferating cells, but also with altered DNA methylation status in cold-stressed quiescent cells. Nucleic Acids Res 28, 3250-3259.


involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. Proc Natl Acad Sci U S A 97, 11632-11637.


<table>
<thead>
<tr>
<th>Main characteristics</th>
<th>TGS in <em>trans</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Target(s)</td>
<td>Transgene and endogene</td>
</tr>
<tr>
<td>Presence of target nuclear RNA?</td>
<td>N/A</td>
</tr>
<tr>
<td>Presence of target cytoplasmic RNA?</td>
<td>No</td>
</tr>
<tr>
<td>Methylation of target(s)</td>
<td>Yes (transgene)</td>
</tr>
<tr>
<td></td>
<td>N/A (endogene)</td>
</tr>
<tr>
<td>Inducing agent(s)</td>
<td>Transgenes expressing promoter dsRNA</td>
</tr>
<tr>
<td></td>
<td>Or the structure of direct repeats</td>
</tr>
<tr>
<td>Transcription of inducer required?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Or not necessary (endogene)</td>
</tr>
<tr>
<td>Production of a systemic signal?</td>
<td>Suggested</td>
</tr>
<tr>
<td>Release by non-homologous viruses?</td>
<td>N/A</td>
</tr>
<tr>
<td>Mutations that release silencing</td>
<td><em>som 8, mom1</em></td>
</tr>
<tr>
<td>Mutants that do not release silencing</td>
<td>N/A</td>
</tr>
<tr>
<td>Silence of transgene required?</td>
<td>Yes (transgene)</td>
</tr>
<tr>
<td></td>
<td>No (endogene)</td>
</tr>
</tbody>
</table>

**Abbreviations:** N/A, not available

Note: updated based on H Vaucheret and M Fagard, 2001 Trends in Genetics Vol. 17, 29-35. Underlined characteristics are added based on this work.

Table 5.1 Update of characteristics of plant TGS in *trans*. 

144
Figure 5.1 *CBF* and *COR* gene expression analyses in methylation mutants.

(A) RNA gel blot analyses of wild type Zürich (Z), line A (L-A), *som* 8 (S-8), *mom* + (M+), and *mom* – (M-). Plants growing at 22 °C were subjected to 4 °C and harvested 4 h and 6 h later. Eight ug of total RNA was loaded per lane. Multiple lanes having the same time point are independent RNA from the same biological samples. The same filter was hybridized, stripped, rehybridized using *CBF1, COR15a, ABF1, HPT*, and *eIF4A* gene probes. *eIF-4A* was used as loading control. (continued)
Figure 5.1 (continued)

(B), (C), (D), and (E) Quantitative representation of relative gene expression shown in A. *CBF1*, *COR15a*, and *ABF1*, hybridization signal bands at 4 or 6 h were quantitated and normalized to *eIF-4A*. The 4 h time point represents the mean of two or three (line A) samples. For *HPT* gene, six signal bands from S-8 line or M + line were all quantified and averaged. The error bar is the standard error of the mean. RNA gel blot was repeated once. Zero time point was not included in the bar figure.
Figure 5.2  Freezing assay of mutant lines.

Seedlings of wild type Zürich (Z), som 8 (S-8), mom + (M+), and mom – (M -) were grown on Gamborg B5 plates at 22 °C for two weeks. Plates were transferred to –2 °C at dark for 3 h and then misted with sterile cold water. Plates were maintained at –2 °C for additional 21 h before decreasing the temperature to –6 °C. After at –6 °C for 24 h, plates were moved to 2 °C and kept in the dark for 8 h, and then at 4 °C for 16 h. Plants were then incubated at 22 °C with 16 h light/8h dark cycle for 20 h before photo documentation.
Figure 5.3 A model for multiple copies of *HPT* gene in one locus epigenetically regulating genes involved in plant cold acclimation.

Three *CBF* genes are physically linked on Arabidopsis chromosome 4 (Gilmour et al., 1998). Multiple copies (5 to 10) of *HPT* gene integrated at a single locus (indicated by a circle). Dotted line indicates the locus is either on the same or different chromosome with CBFs (chromosome 4). Integrated repeats can repress neighbor CBF genes locally by creating a silencing structure (in *cis*) or more likely act on selective sites on different chromosomes (in *trans*). The *trans* model acts either through RNA-directed DNA methylation of promoter region (upper lines, current understanding) or acts through unknown mechanism to selectively regulate the expression of genes involved in cold acclimation (lower lines, suggested by this work). Reduced CBF transcripts results in less CBFs, represented by smaller font, and smaller arrows, leading to less expression of *COR15a*. Note silence triggered by the structure of transgene repeats rather than its status of silence, and only show the CBF genes.
CHAPTER 6

ARABIDOPSIS HOMOLOGS OF TRANSCRIPTION DOMAIN-ASSOCIATED PROTEIN 1 (TRA1) ARE ESSENTIAL FOR EMBRYOGENESIS, GROWTH, AND DEVELOPMENT

6.1 ABSTRACT

Plants need to integrate internal and external cues to regulate gene expression in appropriate time and space by cascades of transcription factors. Yeast Tra1 and its homologs, which also serve as the direct targets for many trans-activators, are critical and only common components of histone acetyltransferase complexes (HATs) involved in chromatin-modifying that facilitate transcriptional activation. Yet whether Arabidopsis has similar protein factor playing such important roles remains unclear. Here we show that Arabidopsis encodes two Tra1-like genes with high identity over their entire length. We designated them as \textit{AtTra1-2} and \textit{AtTra1-4}. Protein sequence analyses revealed that \textit{AtTra1} proteins contained multiple protein-protein interaction motifs with putative function in transcriptional regulation. Homozygous T-DNA insertion mutants in \textit{AtTra1-2} or \textit{AtTra1-4} had no obvious phenotype under normal growth conditions. However, double homozygous T-DNA mutants appeared to be lethal, indicating that AtTra1 was
essential and that two genes might be functionally redundant. Mutants with heterozygosity in one AtTra1 gene and homozygosity in the second AtTra1 gene develop normally, but siliques contain aborted embryos, which indicates that lethality occurs sometime during embryogenesis. RT-PCR analyses suggested that T-DNA insertions resulted in truncated transcripts that were absent the conserved ataxia telangiectasia mutated (ATM) PI-3 kinase (PI3K) domain. Segregation analyses indicated that embryo lethality was associated with mutations in C-terminus containing PI3K domain. RNAi mediated knockdown expression of AtTra1 resulted in pleiotropic phenotypes on plant growth and development, including elongated hypocotyls, smaller and abnormally colored cotyledons. Those phenotypes are similar to those of ada2b-1 mutants, suggesting they possibly function in a same pathway. AtTra1-2 promoter::GUS fusion constructs indicated that AtTra1-2 was expressed in meristem and miotically active tissues. Thus we conclude that AtTra1 genes are essential to Arabidopsis embryogenesis, growth, and development. AtTra1-2 may also serve an important functional role in miotically active cells. Several lines of evidence suggest that AtTra1 plays an essential role as a subunit of a SAGA like complex in Arabidopsis.

6.2 INTRODUCTION

Plants have many more transcription factors (TFs) than animals (Arabidopsis Genome Initiative, 2000). The most recent annotation of the Arabidopsis genome indicates that 6.5% of total 26,207 protein-coding genes encode TFs (Haas et al., 2005). This is thought to be partially due to a sessile nature. Plants must cope with and adapt to
changing environment during their lifetime. Thus, plants need to constantly integrate internal and external cues to regulate gene expression in appropriate time and space. During this regulation, transcription factors (activator or repressor) play a key role.

Transcriptional activators typically contain a DNA binding domain to determine the specificity of regulation and an activation domain (AD). The main role of an AD is believed to recruit protein cofactors for activation. Fusion of a DNA binding domain to a subunit of holoenzyme complex results in the transcription of the target gene without an activation region (Ptashne and Gann, 1997).

Chromatin-remodeling factors associate with TF (Naar et al., 2001). Gene specific activation is a process of ordered recruitment. Activators first bind to cognate regulatory sequences and then recruit chromatin-remodeling factors and RNA polymerase II holoenzyme complex to promoter region for activation (Cosma, 2002). Genetic and biochemical evidence suggests that two different family of chromatin modifying factors (ATP-dependent SWI/SNF complex and histone acetyltransferase complex) are the major targets for activators (Pollard and Peterson, 1998; Fry and Peterson, 2001). SWI/SNF complexes alter DNA-histone contacts by using the energy released from ATP hydrolysis, whereas histone acetyltransferase (HAT) complexes modify histone by acetylation. One well characterized HAT complex is the yeast SAGA (Spt/Ada/GCN5/acetyltransferase) complex, which acetylates histone H3 by subunit GCN5 with the requirement of Ada2 SANT domain (Brown et al., 2000; Sterner et al., 2002). Recently, mammalian TRRAP (transformation/transcription domain associated protein) and its yeast homolog, Tra1 (transcription domain associated protein 1) were identified as the subunits of SAGA
complex and also serve as the direct targets for many transcription factors (McMahon et al., 1998; Saleh et al., 1998; Grant et al., 1998a).

Tra1/TRRAP plays essential roles in many signal transduction pathways. Tra1/TRRAP is the only common component shared by many HAT complexes (Grant et al., 1998a). For example, TRRAP is the subunit of SAGA, NuA4, STAGA, and TFTC complexes (Saleh et al., 1998; Vassilev et al., 1998; Ard et al., 2002). The carboxyl termini of both proteins exhibit a significant homology to the PI 3-kinase (phosphatidylinositol 3-kinase) of ataxia telangiectasia mutated (ATM) family (Keith and Schreiber, 1995; McMahon et al., 1998; Grant et al., 1998a). This PI 3-kinase domain (referred to as PI3K) of TRRAP is required to form a structural core to assemble HAT complexes, which is recruited to promoter region by transcriptional activators (Park et al., 2001). The mutations in ATM related PI3K domain result in the loss of c-Myc mediated oncogenesis. Interestingly, the different regions of C-terminal Tra1/TRRAP, including the PI3K, have protein-protein interactions with many well characterized transcriptional activators as direct targets (Ard et al., 2002). This list includes VP16 (Utley et al., 1998; Brown et al., 2001; Hall and Struhl, 2002; Klein et al., 2003), GAL4 (Bhaumik and Green, 2001; Brown et al., 2001; Klein et al., 2003; Bhaumik et al., 2004), GCN4 (Utley et al., 1998; Brown et al., 2001; Klein et al., 2003), and HAP4 (Brown et al., 2001).

Interestingly, SAGA function predominates in the regulation of stress-related genes in yeast genome. Although the SAGA complex and the TFIID complex make overlapping contribution to regulate the expression of nearly all genes, genome-wide examination indicates that SAGA complex is likely the major contributor to modulate genes upregulated in response to a set of environmental stresses, such as heat, acidity,
oxidation, carbon or nitrogen starvation, and DNA damage (Huisinga and Pugh, 2004). More interestingly, most of those stress-related genes in yeast are TATA box-containing genes (Basehoar et al., 2004). SAGA complex likely involves recruiting TATA-box binding protein (TBP) and basal transcription machinery to these promoters via the interaction with many sequence-specific TFs as described above (Larschan and Winston, 2001; Bhaumik and Green, 2002; Timmers and Tora, 2005).

It is fundamental to determine whether Arabidopsis has similar complex to regulate gene expression, especially the stress-related genes. This regulation is particularly important for plants considering their sessile nature. Arabidopsis transcriptional activator CBF1 (C-repeat bind factor 1) plays a “master switch” role in the regulation of genes contributing to cold tolerance in Arabidopsis and a similar role in crop species (Stockinger et al., 1997; Sarhan, 1998). CBF1 requires yeast ADA2, ADA3 and GCN5 genes, the key components of ADA and SAGA histone acetyltransferase complexes, to be fully functional in trans-activation in yeast (Stockinger et al., 2001). Moreover, CBF1 interacts with Arabidopsis homologous ADA2 and GCN5 in vitro. AtGCN5 also has HAT activity (Stockinger et al., 2001). However, it is the DNA binding domain, AP2 domain, of CBF1 interacting with ADA and GCN5 but not the activation domain (Mao et al., 2006). Our extensive mutagenesis analyses revealed that the multiple hydrophobic motifs in CBF1_AD redundantly contributed to the trans-activation (Wang et al., 2005). This feature is conserved among several other acidic transcriptional activators from other model systems, which directly interact with Tra1, including VP16, GCN4, and Hap4 (Drysdale et al., 1995; Jackson et al., 1996; Sullivan et al., 1998; Stebbins and Triezenberg, 2004). The Arabidopsis genome encodes two homologous Tra1 genes. We
named \textit{AtTra1-2} and \textit{AtTra1-4}. Thus, we hypothesized that the Arabidopsis Tra1 homologs might be the target for CBF1 activation domain. To continue our investigation on the understanding of CBF1 mediated plant \textit{COR} gene trans-activation, we determined the functional roles of two Arabidopsis homologous Tra1 proteins in the regulation of gene expression, growth and development. Here, we present our understanding that the functions of homologous Tra1 protein are essential for Arabidopsis embryogenesis, growth and development.

6.3 MATERIALS AND METHODS

Plant Materials, Growth Conditions, and Stress Treatments

\textit{Arabidopsis thaliana} (ecotype Columbia) was grown on Gamborg’s B5 medium under a 16-h-light/8-h-dark cycle at 22°C for 3 weeks. For cold treatments, the seedlings were placed at 4°C and then harvested at the times indicated in the figure. For freezing assays, the seedlings were first placed at 4°C for two weeks. The temperature was then dropped to -2°C for 24 h, and then to -5°C for 48 h. Temperature was increased to 4°C for 48h to completely thaw the ice and then returned to 22°C for an additional 48 h. The damage caused by freezing temperature was documented.

Identification and Molecular Characterization of Single and Double Mutants

SALK T-DNA insertion mutant lines in the \textit{AtTra1-2} and \textit{AtTra1-4} gene were obtained from the Arabidopsis Biotechnology Research Center (ABRC, Ohio State University, Columbus, OH). These lines were first germinated on Gamborg’s B5 medium supplemented with 30 µg/ml kanamycin, and subsequently transferred to
planting mix. Gene specific primers and primers for the T-DNA left border (Table 6.1) were designed and used to verify the insertions and determine whether lines were homo- or heterozygous. Genotyping PCR reactions were performed by denaturation at 96 °C for 2 min, followed by 35 cycles of incubation at 94 °C for 30 s, 50~60 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 10 min.

Arabidopsis genomic DNA was extracted using the final buffer mixed as follows: 1x DNA extraction buffer (0.35M Sorbitol, 0.1M Tris, 0.005 M EDTA, pH7.5, and 0.02M NaBisulfite was added just before use), 1x nuclei lysis buffer (0.2 M Tris, 0.05 M EDTA, 2.0 M NaCl, 2% CTAB, pH7.5) and 0.4x 5% Sarocosyl were mixed together. One leaflet from each seedling was ground in the 200 µl final buffer. After grinding, an additional 450 µl of the final buffer was added. Ground tissue in the final buffer was incubated for 20-60 min at 65 °C. 600 µl of chloroform was then added and the mixture emulsified. Samples were spun for 7 min at 13,000 rpm. The supernatant was transferred to a new tube. Finally 650 µl isopropanol was added to the supernatant to precipitate genomic DNA for 10 min at 13,000 rpm. Pellets were washed with 70% ethanol once. Dried pellets were dissolved in 200 µl ddH₂O. 2 µl was used for PCR reaction.

Homozygous T-DNA insertion mutants of *AtTra1-2* were obtained from two lines, SALK_050746 and SALK_082716, which were renamed AtTra1-2D and –2G, respectively. For *AtTra1-4*, three homozygous lines were obtained from SALK_087867, SALK_105303, and SALK_053301, and renamed AtTra1-4L, -4K, and -4N, respectively. Double mutants were generated between T-DNA insertion mutants AtTra1-4K and
AtTra1-2D, AtTra1-4N and AtTra1-2D, after the lines were confirmed to be homozygous. The F1 were allowed to self. We then focused on the progenies of the –4K x –2D cross. About 160 F2 seedlings from the –4K x –2D cross were genotyped by PCR.

**Construction of the Constitutive and Inducible RNAi Constructs**

To generate the hairpin RNAi construct, we amplified a 900 bp fragment of the AtTra1-2 PI3K domain (Figure 6.14 A), from the *in vitro, en masse* excised cDNA library CD4-16 (ABRC, Ohio State University) using primers: e993, GCTCTAGAGGCGCGCCATGAACAAATGTTTGATAAG and e994, CGGGGATCCattaaatCATCATAAGTTGCGAG. For cloning purposes, inner and outer restriction sites were incorporated into primers. Primer e993 harbors Asc I (inner, and italic) and Xba I (outer, and underlined) sites; e994, Swa I (inner, and lower case letter) and BamH I (outer, and bolded) sites. This PCR fragment was cloned into the vector pGEM-Teasy (Promega, Madison, WI), confirmed by sequencing, and named pZW117.

Two types of RNAi constructs, a constitutive promoter construct and an inducible promoter construct, were generated. To generate the constitutive promoter construct, we first removed a 2.8 kb EcoR I–Xba I fragment containing the CaMV35S promoter and CHSA intron (Figure 6.14 B) from binary vector pFGC5941 (ABRC, Columbus, OH). This EcoR I–Xba I fragment was subcloned into the EcoR I–Xba I sites of pSP72 (Promega, Madison, WI), creating pZW140. The Asc I–Swa I fragment of pZW117 was then subcloned into the Asc I–Swa I sites of pZW140, creating the sense insert plasmid pZW144. To create the anti-sense insert, the BamH I–Xba I fragment of pZW117 was ligated to the corresponding sites of pZW144, creating pZW145. Next, we cut out the
EcoR I–Xba I fragment, the 35S-Tra1sense-ChsA-Tra1anti cassettes, and cloned back to pFGC5941 binary vector as the constitutive RNAi construct pZW146 (Figure 6.14 B).

To generate a cold inducible RNAi construct, we PCR amplified an 844 bp fragment of the COR15a promoter region using primers: S11, 5′-GGAGATGTTACTGTCCGTCAGA-3′ and S12, 5′-CGC GGATCCATGTTGTTTGAAATGAAAGGAGGA-3′). An Nco I restriction site (underlined) was incorporated into S12 for cloning purposes. The PCR product contains an endogenous EcoR I site resided in the 5′ end of PCR product. The amplified fragment was cloned into the vector pGEM-Teasy, creating pZW148. The 639 bp EcoR I–Nco I fragment of pZW148 was used to replace the CaMV35S promoter of pZW146. The newly resulting plasmid was named pZW169 (Figure 6.14 B).

**Construction of AtTra1 Promoter::β-Glucuronidase Fusions and Glucuronidase Staining**

To generate the AtTra1 promoter::β-glucuronidase fusion constructs, we amplified the upstream promoter regions of AtTra1-2 and AtTra1-4. Presently no Arabidopsis ESTs exist for the predicted 5′ end of AtTra1, however there are poplar (Populus alba x P. tremula var qlandulsa) ESTs that align with the (TAIR annotated) 5′ most predicted exons of AtTra1-2 and AtTra1-4. Poplar cDNA clone PO 02035A02 sequence was obtained using AtTra1-2 sequence as a BLAST query to search the database of GenBank. Alignment of this poplar cDNA with DNA sequences upstream of predicted AtTra1 start codon ATG was performed using ClustalW (http://www.ebi.ac.uk/clustalw/) (Figure 6.12). Then a 665 bp fragment upstream of the predicted ATG start codon of AtTra1-2
(intergenic region between At2g17940 and At2g17930) was PCR amplified from genomic DNA using primers S72 and S74 (Table 6.1). Primer S72 incorporated a *Hind* III site; S73, a *Bgl* II site. The amplified fragment was subcloned into pGEM-Teasy creating pZW126. Plasmid pZW126 was digested with *Hind* III and *Bgl* II, and the *Hind* III–*Bgl* II fragment was subcloned into the *Hind* III–*Bgl* II restricted binary vector pBI121, replacing the 800 bp CaMV35S promoter region. The resulting plasmid was designated pZW180. For AtTra1-4, we PCR amplified a 715 bp fragment of upstream promoter region using primers S75 and S76 (Table 6.1). Primer S75 incorporated a *Hind* III site and S76 a *Bgl* II site. The PCR product was cloned to pGEM-Teasy, sequenced to confirm and designated pZW128. Plasmid pZW128 was restriction digested with *Hind* III and *Bgl* II, and the resulting fragment ligated to the *Hind* III–*Bam*HI sites of pBI121 creating pZW182.

Transgenic plants transformed with pZW180 and pZW182 were identified by the screening with kanamycin 35 µg/ml. Kanamycin resistant T1 seedlings were stained for β-glucuronidase (GUS) activity with 5-bromo-4-chloro-3-indolyl-β−D-glucuronide overnight. The next day seedlings were incubated in 50% ethanol to remove chlorophyll (Jefferson et al., 1987). Due to the problem of kanamycin resistance of vector pBI121, so far I screened four transgenic seedlings from pZW180 transformed seeds. For pZW182, new screening is underway.

**Plant Transformation and Screening of Transgenic Seedlings**

RNAi and promoter::β-glucuronidase fusion binary constructs were electroporated into *Agrobacterium* strain GV3101. Arabidopsis plants (Columbia ecotype) were transformed using the floral dip method (Clough and Bent, 1998). Transgenic T1
seedlings were screened on GB5 media supplemented with either 35 µg/ml kanamycin (pZW180 and pZW182 constructs), or 10-15 µg/ml BASTA (pZW146 and pZW169 constructs).

**RNA Gel Blot Analysis**

Total RNA from T-DNA insertion and RNAi mutant lines was extracted using QIAGEN Plant RNeasy Kit (QIAGEN, Valencia, CA). Filters were hybridized in succession with following probes: *CBF1, ABF1, COR47, COR6.6, COR78*, and *COR 15a*. The RNAs stained by ethidium bromide was used as loading control. DNA probes were radiolabeled with ³²P-nucleotides using the Random-Prime Labeling Kit (Amersham Pharmacia Biotech Inc., Piscataway, NJ). *COR47 and COR6.6*, as well as *COR78 and COR15a*, were labeled and hybridized at the same time. RNA hybridization results were visualized by phosphorimage autoradiography using a Storm 860 PhosphorImager (Amersham Pharmacia Biotech Inc., Piscataway, NJ).

**Reverse Transcriptase (RT)-PCR Analyses**

RT-PCR analyses were carried out either by a two-step RT-PCR or a one-step RT-PCR reaction. For a two-step RT-PCR reaction, total RNA was first treated with DNase I (Ambion, Austin, TX) to remove DNA. DNA-free RNA was then reverse transcribed using SuperScript III RNase H- reverse transcriptase (200-400 units per reaction) and priming with oligo(dT)₂₀ (Invitrogen, Carlsbad, CA) and incubation at 50 °C for 1 h. PCR amplification was performed using 2 µl synthesized first strand cDNA by initial denaturation at 95 °C for 2 min, followed by 35 cycles of denaturation, annealing and elongation using 95 °C for 30 s, 50 °C 40 s, at 72 °C 1-3 min, and a final extension at 72 °C 10 min. For one-step RT-PCR, 1 µg total RNA was incubated with gene specific
primers using the QIAGEN OneStep RT-PCR kit following manufacture’s recommendations (QIAGEN, Valencia, CA). The gene specific primers are listed in table 6.1. One step RT-PCR amplifications were performed by reverse transcription at 50 °C 30 min, then denaturation at 95 °C for 15 min to activate Taq polymerase, followed by 50 cycles of denaturation, annealing and elongation using 94 °C for 40 s, 50 °C 1 min, at 72 °C 3 min (1kb/min), and a final extension at 72 °C 10 min.

**Cloning of AtTra1 cDNA Fragments**

To molecularly isolate a full length AtTra1 cDNA clone, we used two strategies, RT-PCR amplification from total RNA and PCR amplification from *in vitro* excised cDNA (Chapter 4). Gene-specific primers were designed to selected regions such that the overlapping ends shared at least one common unique restriction site for subsequent cloning purposes (Table 6.1, Figure 6.1A). Primers that were successfully used for PCR amplification were also used to amplify cDNA fragments in RT-PCR reactions (Table 6.1, Figure 6.1B). PCR products were separated on 0.8% agarose gel. The size of products was determined by comparison with 1 Kb DNA ladder (New England Biolabs, Ipswich, MA).

**Phenotypic Analyses**

*AtTra1-2* and *AtTra1-4* T-DNA insertion mutants were grown in both the greenhouse and in growth chambers to determine whether there were visual phenotypic effects resulting from the insertions. A defective seed phenotype was noted in the *AtTra1-2/AtTra1-4* double homozygote and then documented under the dissecting
microscope. The number of defective seeds in the AtTra1-2Dx AtTra1-4K F₃ was determined using a dissecting microscope. A minimum of 10 siliques from each plant was opened and seed counted.

**Bioinformatics Tools to Perform Protein Functional Analyses**

AtTra1 protein Bioinformatic analyses were first conducted using InterProScan. InterProScan runs protein sequence analyses by combined different protein signature recognition methods (Zdobnov and Apweiler, 2001). To predict subcellular localization sites, WoLF PSORT was used ([http://wolfpsort.seq.cbrc.jp/](http://wolfpsort.seq.cbrc.jp/)). WoLF PSORT utilizes known motifs, and is also able to use correlative sequence features like amino acid content (Horton, 2006). PROSITE was used to predict ATP binding sites ([http://ca.expasy.org/prosite/prosuser.html](http://ca.expasy.org/prosite/prosuser.html)). Alignments were conducted using ClustalW ([http://www.ebi.ac.uk/clustalw/](http://www.ebi.ac.uk/clustalw/)). Aligned sequences were shaded by BOXSHADE program ([www.ch.embnet.org/software/BOX_form.html](http://www.ch.embnet.org/software/BOX_form.html)).

**6.4 RESULTS**

**Arabidopsis Contains Two Homologous Tra1 Genes with High Identity**

To identify Arabidopsis Tra1 homologs, the yeast Tra1 gene sequence was used as a BLAST query to perform searches of the Arabidopsis genomic sequence (Arabidopsis Genome Initiative, 2000). BLAST results indicated there were two homologous Tra1 genes, located on chromosome 2 (At2g17930) and chromosome 4 (At4g36080), respectively. We designated these genes as *AtTra1-2* and *AtTra1-4*. 
The coding regions of \textit{AtTra1-2} and \textit{AtTra1-4} are predicted to be 11,388 bp and 11,520 bp, respectively (Arabidopsis Genome Initiative, 2000). However, the full-length cDNAs were not reported for either gene. Nor were there ESTs encompassing the 5' end of either gene. To clone a full-length cDNA, we tried to PCR amplify the sections of the predicted cDNAs using our previously \textit{en masse} excised phagemid cDNAs and total RNAs as templates. Six pairs of primers corresponding to annotated exons of \textit{AtTra1} genes were used for RT-PCR (Figure 6.1A). Currently five of the six designed fragments for AtTra1-4, including fragment 1 (F1), F2, F4, F5, and F6, were amplified. F4 contained a 25 amino acid deletion compared with annotated sequences. F5 and F6 were all confirmed, whereas F1 and F2 were partially sequenced (Figure 6.1B). The remaining sequences of F1 and F2 are currently under confirmation with newly synthesized primers. RT-PCR products covering the region between primer S158 and S155 was amplified recently and waiting for confirmation. The goal is to then ligate the six PCR products together in stepwise fusion to generate the full-length cDNA clone. Sequencing of the RT-PCR products revealed that \textit{AtTra1-4} had 34 exons (same as \textit{AtTra1-2}) and not 35 exons predicted by TAIR annotation. The 28 th intron is actually part of 29 th exon of \textit{AtTra1-4}. Our data indicated that \textit{AtTra1-4} (At4g36080) encodes a 3828 amino acid protein. AtTra1-2 and AtTra1-4 share 86% identity to each other at the protein level (Figure 6.1 C). This high identity is along the entire length of two genes (Figure 6.1 C). However, the Arabidopsis Tra1s share relative low identity to homologs from other organisms. AtTra1-2 is 26% identical and 45% similar to human TRRAP; and 25% identical, and 44% similar to yeast Tra1.
Bioinformatic Analyses of AtTra1 Proteins Reveal Multiple Protein-protein Interaction Motifs and A PI 3-Kinase like Domain, Suggesting Diverse Functional Roles.

Arabidopsis AtTra1-2 and -4 potentially encode 3795, and 3828 amino acid proteins, respectively, that are 86% identical and 91% similar. To identify known domains and motifs, we used Bioinformatic tools including InterProscan, WoLF PSORT, PROSITE, TMPRED (Zdobnov and Apweiler, 2001; Hulo et al., 2004; Horton, 2006). Protein sequence analyses using these programs indicated that the AtTra1s contained multiple motifs important for protein-protein interactions, and a motif related to ATM/PI 3-kinase (Figure 6.1 C and 6.2).

Both proteins contain 14 LxxLL motifs (where L is leucine and X is any amino acid) (Figure 6.1 C, 6.2, and 6.3) dispersed throughout the entire sequences. The LxxLL motif (PFAM PF03867) is present in many transcription factors and cofactors, and mediates protein-protein interactions important in transcriptional regulation (Heery et al., 1997; Litterst and Pfitzner, 2002; Plevin et al., 2005). The first L is assigned +1. The –1 residue, important for binding orientation, is generally either a hydrophobic residue or a small, or polar residue (Plevin et al., 2005). AtTra1-2 or -4 protein contains 6 LxxLL motifs with small or polar residue at –1, and 7 motifs with hydrophobic residue at –1 position (Figure 6.3).

InterProscan program identified three tetratricopeptide repeats (TPR, Figure 6.1 C, 6.2). The TPR is a structural motif with ~34 amino acids in length, that adopts a helix-turn-helix arrangement. Three TPRs are the minimal number to form a superhelix for binding (Gough et al., 2001; Main et al., 2003). The last two TPRs at the C-terminal end
conjugate together, separated by 23 amino acids, whereas the additional TPR resides at the N-terminal ends of both proteins. GENE3D predicted that AtTra1-2 and AtTra1-4 formed tetratricopenptide-like helix (InterPro IPR011990) around the regions containing two TPRs at 2900-3072 for AtTra1-2 and 2948-3120 for AtTra1-4, respectively.

AtTra1 is also related to members in the ATM/PI 3-kinase family. AtTra1s share homology in their C-terminal end 262 amino acid residues with the ATM/PI 3-kinase family (PFAM PF00454) (Figure 6.1 C and 6.2). The members of ATM have a less conserved sequence NH3-terminal to the kinase domain (Keith and Schreiber, 1995). However, AtTra1-2 and AtTra1-4 share high homology over its entire length of sequences (Figure 6.1C). The DxxxxN and DFG motifs of the conserved catalytic site of PI3K are missing in both AtTra1s however (Figure 6.1, 6.2). In its extreme C-terminus, there is the conserved FATC domain (FARP, ATM, TRRAP C-terminus)(Bosotti et al., 2000) with 33 amino acids of AtTra1-2 (3763-3795) and AtTra1-4 (3796-3828) (PFAM PF02260) (Figures 6.1C and 6.4). Two completely conserved Cys residues in the FATC domain of kinase FRAP (FKBP12-rapamycin-associated protein, also called TOR) formed disulfide-bonded loop, which was critical for structural and cellular stability (Figure 6.4 lower panel) (Dames et al., 2005). However, the second Cys residues in Tra1 homologs are His residues instead of Cys (Figure 6.4 upper panel) (Dames et al., 2005).

WoLF PSORT program (Horton, 2006) predicted that the AtTra1 proteins were nuclear localized (Figures 6.1C and 6.2). Three predicted nuclear localization signal (NLS) sequences are present in AtTra1-4 at positions 2026 (PSKRVKI), 2431 (PDMRRKF), and 3518 (RRRH). AtTra1-2 was predicted to have three NLSs at positions 1970 (PSKRVKI), 2392 (PGMRRKF), and 3485 (RRRH). AtTra1-2 also has a
bipartite NLS (KKALQILGKLGRNRRF) at 754 as predicted by WoLF PSORT; AtTra1-4 is almost identical at most residues except for the first residue Gly and the seventh residue Val, which correspond to the residue Lys and residue Leu in the AtTra1-2 bipartite NLS, respectively, and as such was not predicted by WoLF PSORT (Figures 6.1C and 6.2).

WoLF PSORT also identified two leucine zippers (PROSITE PDOC00029) (L-x(6)-L-x(6)-L-x(6)-L) at position 727 (LLELSLTLPSLLPYPRL), and 1805 (LQLATLLLKYQSDLVQRKPHEL) in AtTra1-4. In AtTra1-2 these reside at position 671 (LLELCLTLPSLLPYPRL), and 1749 (LQLATLLLKYQSDLVHQHRKEL) (Figures 6.1C and 6.2). One leucine zipper forms an alpha-helix to dimer with a second alpha-helix formed from a second leucine zipper, resulting in coiled coil structure (O'Shea et al., 1989). Leucine zipper containing proteins involve in DNA-dependent transcriptional regulation (Bucher, 1990).

PROSITE also identified an ATP/GTP-binding site motif A (P-loop) in AtTra1-4 at 2756 (ALVDFGKS), and in AtTra1-2 at 2711 (ALVDFGKS) (Figure 6.1C).

Identification and Molecular Characterization of AtTra1 T-DNA insertion lines

To determine the extent to which two AtTra1 genes might affect plant development and gene regulation, we used a reverse genetic strategy. SALK T-DNA insertion lines harboring T-DNA insertions in AtTra1-2 and AtTra1-4 were obtained from the ABRC (Arabidopsis Biotechnology Research Center). PCR screening using primers flanking the insertion site in combination with a T-DNA left border primer (Table 6.1) confirmed the presence of three T-DNA insertion alleles of AtTra1-4, and two T-DNA insertion alleles
of \textit{AtTra1}-2 (Figure 6.5 A and B). Sequencing of the genomic DNA flanking the insertions confirmed that T-DNAs inserted into the exon 6 (E6), E24, and E26 of \textit{AtTra1}-4. We designated these lines \textit{AtTra1}-4L, -4K, -4N, respectively (Figure 6.5 B). \textit{AtTra1}-2 contained T-DNA insertions in E26 and intron 27 (I27), which we designated \textit{AtTra1}-2D, and –2G, respectively (Figure 6.5 A).

To determine whether the T-DNA insertions affected \textit{AtTra1} expression, we employed an RT-PCR analysis (Figure 6.5 C and D). Multiple primer pairs along the length of \textit{AtTra1} (Table 6.1) were used to determine whether the T-DNA insertions resulted in null mutants, or prematurely truncated transcripts. Using gene specific primers downstream of the insertion site, no RT-PCR product was obtained for the \textit{AtTra1}-2D insertion mutant. However, RT-PCR products were obtained using primers upstream of the insertion site (Figure 6.5 C). These results suggested that the T-DNA insertion in E26 of \textit{AtTra1}-2 resulted in a truncated transcript. In contrast, the T-DNA insertion in I27 of \textit{AtTra1}-2 resulted in PCR products when primers either downstream or upstream of T-DNA insertion were used (Figure 6.5 C). However there was reduced PCR product in comparison to the \textit{eIF-4A} control, which suggested that the T-DNA insertion in I27 might not completely knockout \textit{AtTra1}-2 expression and might be a “leaky” allele (Figure 6.5 C). RT-PCR expression analyses of the \textit{AtTra1}-4L, -4K, and –4N mutants indicated that the region downstream of the insertion sites was not expressed (Figure 6.5 D). RT-PCR analysis of \textit{AtTra1}-4L, -4N, and –4K using primers upstream of their insertion sites is currently in progress.
Analyses of *AtTra1*-2 and *AtTra1*-4 T-DNA Insertion Mutants

Although T-DNA insertions in *AtTra1*-2 and *AtTra1*-4 did not eliminate transcripts altogether, all of the insertional mutants in exons failed to express the region of the protein containing the ATM/PI3K domain. The ATM/PI3K domain is essential for Myc-mediated oncogenic transformation and transcriptional activator recruitment (Park et al., 2001). To determine whether the T-DNA insertions into *AtTra1*-2 or *AtTra1*-4 resulted any detectable phenotype, we grew these T-DNA insertion lines in the greenhouse, and in growth chambers under normal growth condition. The homozygous *AtTra1*-2D, *AtTra1*-4L, -4K, and –4N mutant plants were all indistinguishable from wild type in morphology and development. They all appeared “normal”.

To determine whether the *AtTra1* mutants were affected in their response to cold temperature stress, we subjected insertion mutants *AtTra1*-4L, -4K, -4N to cold temperatures and assayed for *COR* and *CBF* expression patterns. Seedlings grown on GB5 plates were transferred to 4°C and harvested at different times. RNA gel blot analysis was then used to monitor *COR* gene expression. RNA gel blot results indicated that disruption of *AtTra1*-4 had no significant effect on plant *COR* gene expression, including *COR47, COR6.6, COR78*, and *COR15a* (Figure 6.6). Seeds of *AtTra1*-2D mutant were limiting at the time these experiments were conducted. RNA gel blot analyses of *AtTra1*-2D mutants are presently in progress.

*AtTra1*-2 and *AtTra1*-4 Are Expressed in Most Tissues Examined

To determine *AtTra1*-2 and *AtTra1*-4 expression patterns in different plant tissues, we used an RT-PCR strategy. Gene specific primers for *AtTra1*-2 and *AtTra1*-4 (Table
6.1) were used to detect individual gene expression pattern. The data indicated that
*AtTra1-4* was expressed in all tissues and organs examined, including roots, rosette
leaves, stems, cauline leaves, flower buds, and siliques (Figure 6.7), whereas *AtTra1-2*
was expressed in roots, rosette leaves, cauline leaves, and siliques but was not detectable
in flower buds and stems (Figure 6.7).

**Disruption of *AtTra1-2* and *AtTra1-4* Results in Embryo Lethality**

We did not observe an obvious phenotype under normal growth conditions with any
of the *AtTra1* insertion mutants, which suggested that *AtTra1-2* and *AtTra1-4* might be
functionally redundant. This was also supported by data indicating that *AtTra1-2* and
*AtTra1-4* were expressed in most of the same tissues (Figure 6.7). Moreover, these two
proteins share high identity (Figure 6.1 C). To test the hypothesis that *AtTra1-2* and
*AtTra1-4* might be functionally redundant, we attempted to make *Tra1-2/Tra1-4* double
homozygous mutants. Crosses between AtTra1-4K and AtTra1-2D were made.
However, screening of 160 F$_2$ seedlings failed to identify any double homozygous
mutants. F$_2$ plants homozygous at one *Tra1* locus, and heterozygous at the other locus
were recovered, however (Figure 6.8). These data suggested that disruption of both
*AtTra1* genes might be embryo lethal. Indeed, opening up the siliques produced on these
F$_2$ plants revealed both healthy seeds, and small defective seeds that were black in color.
Plants homozygous at one locus, but heterozygous at the second locus (d/d K/k or D/d k/k) also developed normally, suggesting that one allele was sufficient for growth (Figure
6.9).
Additional attempts to recover AtTra1-2D/AtTra1-4K double homozygous plants were also unsuccessful. F₃ seeds from KD5-8 (d/d K/k) and KD5-59 (D/d k/k) were sown and then genotyped. Screening 40 KD5-8 and 45 KD5-59 F₃ individuals failed to identify a single double homozygous mutant (Figure 6.10 B and C). In all F₃ individuals, if one gene was homozygous, the second gene was either wild type, or heterozygous (Figure 6.10 B and C). To determine the frequency of defective seeds present, we visually inspected embryos produced in the siliques of D/d k/k and d/d K/k F₃ genotypes. Well-developed green embryos alongside white and brown embryos were clearly present in siliques of these F₃ plants (Figure 6.11). In contrast, wild type plants, and the single gene homozygous mutants contained only green, normally developed embryos (Figure 6.11). Additional developmental characterization revealed that defective embryos were white at first and latter turned to a purplish/brown (Figure 6.11). White and brown embryos within one silique suggested that embryo development might arrest at different developmental stage (Figure 6.11).

Approximately 20% of the embryos produced on the selfed F₃ D/d k/k or d/d K/k plants failed to develop. In contrast, less than 2% of the embryos failed to develop in selfed WT, or the single AtTra1 insertional mutants (Figure 6.11 A). Thus, we conclude that at least one functional copy of an AtTra1 is essential for Arabidopsis embryo development.

**AtTra1-2 Is Expressed in Meristems and Miotically Active Tissues**

To determine the spatial expression pattern of AtTra1-2 and AtTra1-4, we generated AtTra1-2 and AtTra1-4: β- gulcuronidase fusions. We amplified genomic DNA
fragments upstream of the predicted *AtTra1* ATG start codon and fused these to β-gulcuronidase (GUS) (Figure 6.12). GUS assay of transgenic seedlings showed that *AtTra1*-2:GUS was primarily expressed in meristems (Figure 6.13 A, D, and C) and mitotically active tissues. Young leaves and young floral buds showed intense staining (Figure 6.13 B, C, D and E). GUS staining of the transgenic seedling of *AtTra1*-4: β-gulcuronidase fusion is underway.

**RNAi Reduced *AtTra1* Expression and Results Multiple Growth Phenotypes**

Our attempts to generate *Tra1-2/Tra1-4* double homozygous mutants indicated that at least one functional allele at either of the two loci was essential for embryo development. To further explore the role of *AtTra1* in growth and development, and the plant’s ability to respond to environmental stress, we attempted to reduce the *AtTra1* levels, but not entirely eliminate expression altogether, with an RNAi strategy. RNAi constructs were generated to the PI3K domain (Figure 6.14 A and B). This region is critical for HAT complex assembly and recruitment by transcriptional activators (McMahon et al., 1998; Herceg et al., 2001; Park et al., 2001). Moreover, all of our T-DNA insertion mutants eliminated the COOH-terminal PI3K domain, but none were null allele because the 5′ regions of the genes were still expressed. These data suggested that the PI3K domain might be the essential region for viability. *AtTra1*-2 and *AtTra1*-4 are 86% identical in their nucleotide coding sequences in this region; as such we reasoned that a single RNAi construct might affect RNA levels of both genes. RNAi constructs were placed under the control of the constitutive CaMV35S RNA promoter, and the cold inducible COR15a promoter (Figure 6.14 B).
All T<sub>1</sub> seedlings screened on medium supplemented with BASTA showed varying degrees of elongated hypocotyls, greenness of cotyledon color and morphology (Figure 6.15 A). The hypocotyls of 35S RNAi lines were an average of 6 mm in length, approximately two fold the length of the wild type (Figure 6.15 A and C). RT-PCR analyses indicated that the expression of PI3K domain used to generate the RNAi construct was reduced. In addition, the region upstream of the PI3K domain was also similarly reduced. Taken together, these data suggest that the entire \textit{AtTra1}-2 and \textit{AtTra1}-4 transcripts were effectively reduced (Figure 6.15 B). RNAi lines, 146-2 and 146-6, also showed a dwarf, late flowering, and slower root growth phenotype (Figure 6.15 D).

**Characterization of RNAi Lines under Environmental Stresses**

To determine whether the reduced expression of \textit{AtTra1} might affect the ability of \textit{Arabidopsis} to cold acclimate, wild type and \textit{Tra1} RNAi \textit{T<sub>1</sub>} or \textit{T<sub>2</sub>} seedlings were placed at 4°C for various times. RNA gel blot analyses indicated that \textit{COR} genes in RNAi lines were not significantly affected compared to the wild type control, or seedlings transformed with vector only (Figure 6.16 A, B, and C). Interestingly, in the CaMV35S-\textit{Tra1} RNAi lines, RNAs that cross hybridized with the PI3K probe were present in the warm sample but were almost completely absent in the cold treated plants (Figure 6.16B). Plant freezing assay did not show any significant difference (Data not shown).

To avoid the possible inhibition of RNAi effect under low temperatures (Szittya et al., 2003), we also placed T<sub>3</sub> seedlings of CaMV35S-\textit{Tra1} RNAi lines under drought
conditions. The results were not consistent. Currently, the characterization of these RNAi lines with T1 seedlings under drought condition is in progress.

6.5 DISCUSSION

Arabidopsis Contains Homologous AtTra1

Research in yeast and mammals identified Tra1 or TRRAP, as the only common subunit of several histone acetyltransferase complexes, plays critical roles in regulation of gene expression (Huisinga and Pugh, 2004) and embryogenesis (Herceg et al., 2001). Tra1 or TRRAP plays its essential role by directly interacting with transcription factors. To determine whether Arabidopsis has such important factor with similar critical functional roles, we identified two homologous genes, named as AtTra1-2 and AtTra1-4. The current sequencing results of our RT-PCR products indicate that both genes might contain 34 exons. The 28th intron in AtTra1-4 annotated by TAIR is actually part of exon 29. These two genes share high identity both at the protein level and nucleotide level (Figure 6.1 C and not shown). These two genes might be resulted from Arabidopsis genome duplication (Simillion et al., 2002).

Multiple Protein-protein Interaction Motifs of AtTra1s Suggest Multiple Functional Roles in Multiple Signal Transduction Pathways

Protein sequence analyses revealed that AtTra1s contained multiple protein-protein interaction motifs, including LxxLL motifs, TPR motifs, leucine zipper. LxxLL motifs are short and seemingly simple but have important regulatory roles. LxxLL motifs are
found in many transcription factors and cofactors. Through these motifs mediated specific protein-protein interactions, transcription factors recruit cofactors to promoter region to transcriptionally regulate gene expression (Heery et al., 1997; Litterst and Pfitzner, 2002; Plevin et al., 2005). Motifs with hydrophobic character in the –1 position tend to bind to nuclear receptors as revealed in mammalian system, whereas motifs with small or polar amino acid residues bind to non-nuclear receptors, including transcription factors and mediator subunits (Plevin et al., 2005). AtTra1 proteins have both types of LxxLL motifs (Figure 6.3), which suggest diverse factors interacting with AtTra1. Intriguingly, AtTra1 proteins contain much more LxxLL motifs than its homologs (Figure 6.3). For examples, TRRAP has eight LxxLL motifs (McMahon et al., 1998); yeast Tra1, eight (this work); Drosophila melanogaster homolog, seven (this work); Caenorhabditis elegans, five (this work). C. elegans homolog even only contains motifs with hydrophobic residues but no motif with small or polar residue at –1 position. Given the fact that Tra1, as critical subunit of SAGA complex, plays important role in the regulation of stress-related gene expression in yeast (Huisinga and Pugh, 2004), many more LxxLL motifs than its homologs in Arabidopsis Tra1s may be an important aspect for Arabidopsis to survive under lifelong changing environments. The sessile nature of plants requires it to cope with and adapt to harsh environments by more transcription factor mediated signal pathways. Indeed, Arabidopsis contains more transcription factors than other organism and many of them are unique to plants (Arabidopsis Genome Initiative, 2000). Thus, more LxxLL motifs presumably provide more specific interfaces to bind different transcription factors or cofactors. Through this kind of short peptide-recognition motifs, plants could have economic but efficient mechanisms to exert its
specific roles by transcriptional regulation (activation or repression) in given signal pathways.

In contrast to the short peptide-recognition LxxLL motifs, AtTra1 proteins contain three TPR motifs (Figures 6.1 C and 6.2). The TPR motifs were first identified from genes involved in cell division cycle (Sikorski et al., 1990). The TPR motif is a degenerate sequence with about 34 amino acid residues consisting of two anti-parallel α-helix separated by a turn. Three TPR motifs, as the minimum and most common number in a protein, pack together in a parallel array to form a right-handed superhelical architecture. This arrangement creates a groove, or cradle, to provide a large amount of surface area for ligand binding (Das et al., 1998; D'Andrea and Regan, 2003; Main et al., 2003). Proteins with the TPR motifs usually function as scaffolds for the assembly of multisubunit complexes (Scheufler et al., 2000; D'Andrea and Regan, 2003). One example is from the recent determination of Arabidopsis pale yellow green 7 protein (PYG7), which has three TPR motifs. This small protein (only 296 amino acids) provides basis to assemble Photosystem I (PSI) complex. Pyg7 mutant failed the assembly of PSI but each subunit of PSI is produced (Stockel et al., 2006). The C-terminal two TPR motifs in two AtTra1 proteins are highly identical to each other (Figure 6.1 C). However, the N-terminal TPR motif is not even predicted in the similar position within AtTra1s (Figure 6.1 C). This might be the residue difference in two sequences. Gene3D program predicted that AtTra1-2 forms a Tetratricopeptide-like helical from amino acid 2900-3072; AtTra1-4, 2973-3145. Proteins with the TPR motifs play roles in a wide arrange of biological processes, including photosynthesis, cell cycle regulation,
transcriptional control, mitochondrial and peroxisomal protein transport, neurogenesis and protein folding (Das et al., 1998; D'Andrea and Regan, 2003; Stockel et al., 2006).

WoLF PSORT program predicts AtTra1s are nuclear localized proteins. Interestingly, characterization suggests that each NLS resides adjacent to important protein-protein interaction motifs. For examples, AtTra1-2 bipartite NLS (754) is near three LxxLL motifs (656, 857, and 1052), and leucine zipper motif (672); NLS (1970), near four LxxLL motifs (1357, 1411, 2066, and 2129), leucine zipper motif (1749); NLS (2392), near three LxxLL motifs (2378, 2436, and 2523); NLS (3485), near two LxxLL motifs (2629 and 3264), two TPR motifs (2929 and 3036), and the PI-3K domain (Figure 6.2). Thus, these NLS in AtTra1s might make sure these important protein-protein interaction motifs residing within nucleus.

Multiple protein-protein interaction motifs in AtTra1 explain its multiple interaction sites with many transcriptional activators. Tra1/TRRAP is the direct targets for many transcriptional activators, including GCN4, VP16, and GAL4. However, Tra1/TRRAP interacts with different parts of C-terminal end of Tra1 proteins (Figure 6.14 A) (Ard et al., 2002). This interesting phenomenon might be explained by these multiple LxxLL motifs, TPR motifs in the C-terminal end of Tra1 proteins. LxxLL motifs provide multiple short peptide-recognition motifs, whereas TPR motifs provide a groove, or cradle with a large amount of surface area for ligand binding.

In a sum, protein sequence analyses reveal that AtTra1 contains multiple protein-protein interaction motifs with putative roles in transcriptional regulation. These diverse protein-protein interaction motifs, PI3K, and FATC domain reinforce the expected
multiple functional roles of AtTra1 as a subunit of a putative Arabidopsis complex, with PI 3-K and FATC acting as presumably the catalytic domain.

**AtTra1s Are Essential Implying the Critical Role of Histone Acetylation in Arabidopsis Embryogenesis**

Our results indicated that the disruption of AtTra1 functions results in embryo lethality. This result is consistent with investigations in mice homolog TRRAP (Herceg et al., 2001) and yeast essential Tra1 (Saleh et al., 1998). Mice only contain one TRRAP gene and Trrap heterozygous mutant mice (Trrap$^{+/+}$) showed no defect in development (Herceg et al., 2001). This is also consistent with result of Arabidopsis AtTra1 mutants. As long as Arabidopsis has one wild copy AtTra1 allele, Arabidopsis does not exhibit any obvious phenotype (Figure 6.9).

Tra1/TRRAP has been shown to be critical components of multiple histone acetyltransferase complexes (Grant et al., 1998a; Ard et al., 2002). Our previous work suggested that there were SAGA-like complex in Arabidopsis (Stockinger et al., 2001; Vlachonasios et al., 2003). The characteristic nature of multiple protein-protein interaction motifs of AtTra1s, especially the TPR motifs, suggests it a critical subunit of a multiprotein complex (Figures 6.1 C, 6.2, and 6.3) (Stockel et al., 2006). Thus it is reasonable to consider that Arabidopsis contains similar histone acetyltransferase complexes with AtTra1 as a subunit.

The disruption of AtTra1 genes resulted in embryo lethality. T-DNA insertions eliminated the C-terminal end of AtTra1 transcripts, about 2.75 kb off AtTra1-2 (AtTra1–2D) and 2.537 kb off AtTra1-4 (AtTra1-4N) (Figure 6.5). Thus both T-DNA insertions
eliminate several domains within the C-terminal ends, including two TPR motifs, one LxxLL motif, PI3K, and the FATC domain (Figure 6.2). Research indicated that PI3K of TRRAP was necessary for the histone acetyltransferase recruitment (Park et al., 2001). However, only expression of the C-terminal 428 amino acids of TRRAP was not able to recruit hGCN5 or show the histone acetyltransferase activity. Thus the PI3K itself of TRRAP is not enough to assemble a HAT complex (Park et al., 2001). Furthermore, the expression of TRRAP amino acid 1261-1579 had a significant inhibitory effect (92%) on the c-Myc-mediated oncogenic transformation than the dominant inhibitory effect of the C-terminal 428 amino acids (PI3K and FATC domains) (52%) (McMahon et al., 1998), suggesting an important role of other accessory domains beyond the C-terminal PI3K domain. GENE3D analysis predicted AtTra1 proteins adopting a tetratricopeptide-like helical around the C-terminal two TPR motifs. T-DNA insertions eliminate these two TPR motifs. A small protein pyg7 with three TPR motifs is necessary to assemble PSI complex (Stockel et al., 2006). Thus we think that the TPR motifs function to assemble a HAT complex and act an essential role together with the C-terminal region, including PI3K and FATC domain.

Recent analysis of FATC domain of yeast target of rapamycin (TOR, also known as FRAP) proteins revealed a critical role for TOR structure and cellular stability played by redox potential of a disulfide bridge, which was formed between two completely conserved Cys residues in TOR homologs (Figure 6.4 lower panel) (Dames et al., 2005). Analyses suggested a redox-regulation of the disulfide bond, resulting FATC domain conformation changes and affecting the TOR protein stability (Dames et al., 2005). However, these Cys residues are not completely conserved in Tra1/TRRAP homologs,
which have conserved His residues instead of the second Cys residues (Figure 6.4 upper panel). Thus it is not able to form a disulfide bridge, suggesting that the FATC domains of Tra1/TRRAP might function differently from TOR proteins. To analyze AtTra1 FATC domain, genetic crossing between AtTra1-4K and AtTra1-2F (SALK 087015), a putative T-DNA insertion line with the elimination of the C-terminal end 36 amino acids of AtTra1-2, is in progress. Together, it appears that the accessory protein-protein interaction motifs provide additional help to the necessary PI3K of Tra1/TRRAP to the assembly of functional HAT complexes, which might be essential to Arabidopsis embryogenesis.

Recent work indicated that DNA methylation is also critical to regulate gene expression involved in embryogenesis (Xiao et al., 2006). Disruption of DNA methyltransferases affects the cell division and results in embryo lethality. Our analyses of AtTra1 also imply a critical role of histone acetylation during embryogenesis. Interestingly, a nuclear protein GLUTAMINE-RICH PROTEIN23 (GRP23) with pentatricopeptide repeat (PPR), a TPR-related motif, is also critical to early embryo development (Small and Peeters, 2000; Ding et al., 2006).

**AtTra1s Play Roles in Cell Division, Growth and Development**

It has been established that Arabidopsis embryogenesis is a process with programmed cell division, differentiation (Berleth, 2002). Furthermore, AtTra1 motifs like TPR domains found in cell division cycle genes imply a role in cell division (Sikorski et al., 1990). The results of GUS assay revealed that AtTra1-2 was primarily expressed in mitotically active tissues, suggesting a role in cell division (Figure 6.13).
Knockdown the expression of *AtTra1* genes resulted in pleiotropic phenotypes, including elongated hypocotyls, abnormal cotyledons, and dwarf stature that is partially overlapping with phenotype of *Atada2b-1* mutants (Figure 6.14) (Vlachonasios et al., 2003). Ada is subunits of histone acetylation complexes such as ADA and SAGA complex (Grant et al., 1998b). These overlapping phenotypes suggest that Arabidopsis Tra1 and ADA might function in a same pathway.

**AtTra1 and Environmental Stresses**

Yeast SAGA complex is the major contributor to regulate genes in response to a set of stresses (Huisinga and Pugh, 2004). Previous work and current investigation suggested that Arabidopsis contained the plant SAGA-like complexes (Stockinger et al., 2001; Mao et al., 2006). The DNA binding domain of transcriptional activator CBF1 interacts with Arabidopsis ADA2 and ADA2 interacts with GCN5 (Stockinger et al., 2001; Mao et al., 2006). Furthermore, the expression of plant *COR* genes in homozygous T-DNA insertion lines of *ADA2b* and *GCN5* was reduced but not abolished (Vlachonasios et al., 2003). To determine the role of AtTra1 as a putative direct target for CBF1 to regulate *COR* genes induced in cold temperature stress, we first used the T-DNA insertion mutants. However, the essential nature of *AtTra1* and the functional redundancy made it impossible to test the ability of double homozygous *AtTra1* lines in cold tolerance (Figure 6.11). It appeared that one wild type copy of *AtTra1* allele ensured normal development and the regulation of plant *COR* genes (Figures 6.6 and 6.9). Thus we further tried the RNAi strategy to knockdown the expression of the essential *AtTra1* (Figure 6.14 A and B). RNA gel blot analyses of these RNAi lines did not show
significant effect on the COR gene expression (Figure 6.16). However, we cannot avoid the problem of the inhibition of RNA silencing by low temperatures. This inhibition was observed in plants, insects, and mammalian cells (Fortier and Belote, 2000; Szittya et al., 2003; Kameda et al., 2004). When temperature dropped from 24°C to 15°C, both the virus- and transgene (sense and antisense)-induced RNA silencing were inhibited by the prevention of short interfering (si) RNA generation (Szittya et al., 2003). Further investigation from the same group revealed that the RNA silencing pathway in nine antisense potato lines from 24 total were not inhibited by low temperature. The RNA silencing efficiency might have organ specificity (Sos-Hegedus et al., 2005).

Furthermore, RNA gel blot analyses with transgenic Arabidopsis seedlings expressing the hairpin CBF1 construct showed the inhibition of RNAi effect, suggesting that the inverted repeat induced silence pathway might be affected (Z. W. and E.J.S., unpublished data).

To avoid the low temperature inhibition of RNAi pathway, we tested these RNAi lines under drought condition. Drought and high salinity also induce the expression of plant COR genes partially though CBF4 (Haake et al., 2002). Due to the limited number of T₀ seeds, we used T₃ transgenic seedlings screened on selective medium, which were transferred to soil and withheld the watering for two weeks. However, the phenotypes of different replicates are not consistent. This might be due to the losing of RNAi effect after generations (Mitsuhara et al., 2002). We also observed that only about 10 to 30% seedlings remained elongated hypocotyls within T₃ seedlings in contrast to more than 90% in T₁ seedlings, which was consistent with the results using an intron as spacer to
generate hairpin RNA construct (Smith et al., 2000). Currently using T₁ seedlings to test the COR gene induction under drought condition is in progress.

6.6 ACKNOWLEDGMENTS

We thank the Arabidopsis Biotechnology Research Center (ABRC, The Ohio State University, Columbus, OH) for providing seeds of Arabidopsis T-DNA insertion lines. We thank Erica Jackson for helping with screening of T-DNA insertion lines, and thank Dr David Mackey’s suggestion on embryo lethality. We are also grateful to Dr Sophien Kamoun for using his Leica MZFL110 Fluorescent microscope. This work was supported by Ohio State University, Horticulture and Crop Science Department to E.J.S., and an OARDC Graduate Student Competition Grant to Z. W.

6.7 REFERENCES


encodes a novel nuclear PPR motif protein that interacts with RNA polymerase II subunit III. Plant Cell 18, 815-830.


Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., 
critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. 

acid motif in CDC23 defines a family of proteins and a new relationship among 

Simillion, C., Vandepoele, K., Van Montagu, M.C., Zabeau, M., and Van de Peer, Y. 
(2002). The hidden duplication past of Arabidopsis thaliana. Proc Natl Acad Sci 
U S A 99, 13627-13632.

Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in 


Active RNA silencing at low temperature indicates distinct pathways for 

Stebbins, J.L., and Triezenberg, S.J. (2004). Identification, mutational analysis, and 
coactivator requirements of two distinct transcriptional activation domains of the 

SANT domain of Ada2 is required for normal acetylation of histones by the yeast 

conserved tetratrico Peptide repeat protein pale yellow green7 is required for 
photosystem I accumulation in Arabidopsis and copurifies with the complex. 
Plant Physiol 141, 870-878.

CBF1 encodes an AP2 domain-containing transcriptional activator that binds to 
the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates 
transcription in response to low temperature and water deficit. Proc Natl Acad Sci 
U S A 94, 1035-1040.

Stockinger, E.J., Mao, Y., Regier, M.K., Triezenberg, S.J., and Thomashow, M.F. 
(2001). Transcriptional adaptor and histone acetyltransferase proteins in

186


<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’ to 3’)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>e916</td>
<td>TGGTTCACTAGTGGGCCATCG</td>
<td>T-DNA left border</td>
</tr>
<tr>
<td>e920</td>
<td>TTTTTGAATCGGCACAAAGTC</td>
<td>Fd specific for –2D</td>
</tr>
<tr>
<td>e921</td>
<td>TGTTGTCCATGAAACTTGGCA</td>
<td>Rs specific for –2D</td>
</tr>
<tr>
<td>e936</td>
<td>CACACCCAAAGACAGAAGATGGG</td>
<td>Rs specific for –2D</td>
</tr>
<tr>
<td>e937</td>
<td>AAGGCAAACCGACCAGCATTAT</td>
<td>Rs specific for –4K</td>
</tr>
<tr>
<td>e993</td>
<td>GCTCTAGAGGCACCGCCATGAACCAAATTTTGATATAG</td>
<td>For RNAi construct</td>
</tr>
<tr>
<td>e994</td>
<td>CGGGATCCATATACATCATACATAAGTTGCGAG</td>
<td>For RNAi construct</td>
</tr>
<tr>
<td>ES108</td>
<td>CCGGAATTCGCAGTTGAACCAAGGATGCCATGTTG</td>
<td>Rs specific for –2</td>
</tr>
<tr>
<td>S11</td>
<td>GGAGATGTACTGTCGTCGAGA</td>
<td>For COR15a promoter</td>
</tr>
<tr>
<td>S12</td>
<td>CGCAGATCCATGTTGATTTGGGAATGAAAGGAGGA</td>
<td>For COR15a promoter</td>
</tr>
<tr>
<td>S14</td>
<td>CGGGATCCGGCCGCACCTGATAAGTTGCGAG</td>
<td>Rs specific for –2</td>
</tr>
<tr>
<td>S17</td>
<td>AGTTTAGGTTCCATATAAAAA</td>
<td>Rs specific for –2</td>
</tr>
<tr>
<td>S18</td>
<td>GTAATTTACACATGAAACTTCCAAT</td>
<td>Fd specific for –4</td>
</tr>
<tr>
<td>S20</td>
<td>TCCAGAGTCTCTCTAGAAATTATCG</td>
<td>Fd specific for –4</td>
</tr>
<tr>
<td>S21</td>
<td>GATTTGAGTTTCTCTAGAAATTATCG</td>
<td>Rs specific for –4</td>
</tr>
<tr>
<td>S60</td>
<td>CACAAACATCATATAAGCCCTTG</td>
<td>Rs specific for –2</td>
</tr>
<tr>
<td>S64</td>
<td>CCTCAGCTCCCGATACTCAAT</td>
<td>Rs specific for –4</td>
</tr>
<tr>
<td>S65</td>
<td>GACAAAGCCAGCGGATAGAG</td>
<td>Fd specific for –4</td>
</tr>
<tr>
<td>S72</td>
<td>CAAGCTCTCCTCCTCTCATTACTATCTG</td>
<td>Fd for –2 promoter</td>
</tr>
<tr>
<td>S74</td>
<td>CTCTAGATCTTTGGTAGACAAAGAGCTGAGC</td>
<td>Rs for -2 promoter</td>
</tr>
<tr>
<td>S75</td>
<td>CCAAGCTTTTGATTTTACACAGAAACAAAATA</td>
<td>Fd for –4 promoter</td>
</tr>
<tr>
<td>S76</td>
<td>CTCTAGATCTCATAAAGCTGACCCAGCACAAGC</td>
<td>Rs for –4 promoter</td>
</tr>
<tr>
<td>S118</td>
<td>CCAATTCGTCCTCGTGAATCTACCC</td>
<td>Rs specific for –4</td>
</tr>
<tr>
<td>S120</td>
<td>ACGATTCCAATGGGATGATCGAC</td>
<td>Fd specific for –4</td>
</tr>
<tr>
<td>S121</td>
<td>CGTGGTCCTTGCCTGGATAGGCTTC</td>
<td>Fd specific for –4</td>
</tr>
<tr>
<td>S122</td>
<td>TGCTATGTCTTCTGCGATAGTGAATTTG</td>
<td>Fd specific for –2</td>
</tr>
<tr>
<td>S128</td>
<td>CAGCTTATCCATCTCCTTGGTAAA</td>
<td>Primers for both</td>
</tr>
<tr>
<td>S129</td>
<td>ACACCCAGCTCTCGTCTAAGACTTA</td>
<td>Fd specific for –2</td>
</tr>
<tr>
<td>S130</td>
<td>AAAGTCCAGCTCTTTGGTTGCACGC</td>
<td>Rs specific for –2</td>
</tr>
<tr>
<td>S131</td>
<td>TCACCAACCACCTGTGGATATGCGG</td>
<td>Fd specific for –4</td>
</tr>
<tr>
<td>S132</td>
<td>TTGAAAATCAGCTCTCGATTTGGAATTTG</td>
<td>Rs specific for –4</td>
</tr>
<tr>
<td>S133</td>
<td>CATGTTGACATCCTGCAATGGGAAC</td>
<td>Fd specific for –2</td>
</tr>
<tr>
<td>S134</td>
<td>AGATTAAGACCGACTAGCGCAGCTCTCC</td>
<td>Rs specific for –2</td>
</tr>
</tbody>
</table>

Table 6.1 Oligonucleotides used as PCR primers in the text.

Fd, forward; Rs, reverse; -2, AtTra1-2; -4, AtTra1-4. (continued)
<table>
<thead>
<tr>
<th></th>
<th>Primer Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S140</td>
<td>AGCCTTCTGCTATTCAGCAA</td>
<td>Fd <em>eIF4A</em> gene specific</td>
</tr>
<tr>
<td>S141</td>
<td>TTGTGATCTCCAAAGCTTCC</td>
<td>Rs <em>eIF4A</em> gene specific</td>
</tr>
<tr>
<td>S154</td>
<td>CCATGAGTCCAAATTCAGAATTT</td>
<td>5' end primer for both</td>
</tr>
<tr>
<td>S155</td>
<td>TGCCCATGATCATCTGGTTCAGGCA</td>
<td>Fd primer w/S132 for -4</td>
</tr>
<tr>
<td>S156</td>
<td>AGCGTACTAGCTCCATCCTG</td>
<td>Rs primer w/S157 for -4</td>
</tr>
<tr>
<td>S157</td>
<td>GATCAACAGGGTAAGACTGTGA</td>
<td>Fd primer w/S156 for -4</td>
</tr>
<tr>
<td>S158</td>
<td>AAGCTGCTGCACTCAGTATGG</td>
<td>Rs primer w/S159 for -4</td>
</tr>
<tr>
<td>S159</td>
<td>TGTCGAGCTGCTTTTCTAGGGATTTAG</td>
<td>Fd primer w/S158 for -4</td>
</tr>
<tr>
<td>S160</td>
<td>CCCTAAGTCAGCCTATTTTCGT</td>
<td>Rs primer w/S161 for -4</td>
</tr>
<tr>
<td>S161</td>
<td>CTGGGTACGCTTTTTTTATGCA</td>
<td>Fd primer w/S160 for -4</td>
</tr>
<tr>
<td>S162</td>
<td>GGATCTTGACCCTACCTATTCACAC</td>
<td>Fd specific for -2</td>
</tr>
<tr>
<td>S163</td>
<td>AGCTGACCTACCTATTGAGGA</td>
<td>Fd specific for -4</td>
</tr>
<tr>
<td>S164</td>
<td>CGAGAAGGGACAAATCTGCT</td>
<td>Rs specific for -4</td>
</tr>
<tr>
<td>S165</td>
<td>AAAGTGACCAGATCTTTCCAC</td>
<td>Rs specific for -2</td>
</tr>
</tbody>
</table>
Figure 6.1 Arabidopsis contains two *AtTra1* genes.

(A) Schematic representation of *AtTra1*-4 and the six RT-PCR fragments used to generate a full-length cDNA clone (Forward primer, rightward arrow; reverse primer, leftward arrow). Overlapping fragments share at least one unique restriction site for subsequent cloning.

(B) RT-PCR products. (1), RT-PCR products of the 5′ gene ends. Primer S154 starts from the initiation codon ATG. Primers S162 and S165, at similar position with S163 and S164 in homologous *AtTra1*-2, were used to amplify the homologous region of *AtTra1*-2. (2), Primer pairs S161/S160 were used to amplify fragment 1; S159/S158, fragment 2; S157/S156, fragment 3; S155/S132, fragment 4. (3), Primer pairs S65/S21 were used to amplify fragment 5. Primers S64, S60, S128 were paired with S65 for RT-PCR. (4), S20 and S118 were used to amplify fragment 6. Primer pairs S18/ES108, S18/S118, and S18/S14 were used to amplify the C-terminal end of *AtTra1*-4. Markers are 1 kb DNA ladder.

(C) Alignment of the predicted *AtTra1*-2 and *AtTra1*-4 proteins. *AtTra1*-2 and *AtTra1*-4 are the translated polypeptides, determined in part from the annotated Arabidopsis genomic sequence, and in part from the cloned cDNAs. Multiple LxxLL motifs are pink underlined letters in both proteins. Blue characters identify the three predicted nuclear localization signals (NLS); the overlined blue text identifies the bipartite NLS presenting at *AtTra1*-2. The two leucine zipper patterns are italicized green. TPR motifs are in red. Amino acids highlighted in brown identify the ATM/PI3K domain. The PI3K catalytic sites are indicated in dark yellow. Note however that both Arabidopsis proteins have a Gly residue instead of the critical Asp residue in kinase (DxxxN), and a His residue instead of Gly residue (DFG). The 33 amino acid FATC domain of the C-terminus (turquoise) is FRAP, ATM, TRRAP C-terminal (Bosotti et al., 2000). The following sequences of *AtTra1*-2 was confirmed by sequencing RT-PCR products or EST sequences, including 8671-9510,10121-11396; *AtTra1*-4, including 30-1059, 1779-2815, 3695-4694, 5268 to stop codon.
AtTra1-2  MSPIQNFEQHSSRLVDDLPIPTRLEMVVEVRDSLEIAHTAEYLNFLKCYFPFAFSVILLQ  60
AtTra1-4  MSPIQNFQEOHSSRLVEPDPIEERLAMVVEVRDSLEIHTAELYNFLKCYFRASSVILLQ  60

AtTra1-2  ITKPQFIIDNPEHKLRNIVVEILNRLPHSEVLRPFVQDLLKVAMQVLTADNEENGLICIRI  120
AtTra1-4  ITKPQFTDNIEHKLRNIVVEILNRLPHSEVLRPFVQDLLKVAMQVLTADNEENGLICIRI  120

AtTra1-2  IFDLLRNFRPTLENEQVFPLDFVCKIYISIFRFVTSHFFDNVKMEEVKPEMEMPTSSDQLT  180
AtTra1-4  IFDLLRNFRPTLENEQVFPLDFVCKIYQNFRLTVSHFFENVKMEEVKPIETPDQLSLS  180

AtTra1-2  PTPPQIYNPSTRSFKIIITESPLVVMFLFQLYSRLVQNPHLLPLMVAAISVPGEF  240
AtTra1-4  ITAPSNQINPSTRSFKIVTESPLVMFLQYSLRYVQINPNLLPLMVAAISIPGEK  240

AtTra1-2  VPSHLPQFIELGAQVRK--------------------------------------------------------------------------------------------------  258
AtTra1-4  VSSHMKPQFIELGAQVRSTSPFLTYLKSRAEYIJKPEESICKSIVNLLTVCDSASIRK  300

191
AtTra1-2 is 3795 amino acid residues, and AtTra1-4 is 3828 amino acid residues. The 14 LxxLL motifs (x standing for any residue) implicated in protein-protein interaction, are indicated in gray horizontal lines. The 34 amino acid TPR (tetratricopeptide repeat) implicated in ligand binding are indicated in red horizontal lines. Two leucine zippers are indicated in green horizontal lines. Three NLS (nuclear localization signal) and one bipartite are indicated in blue horizontal lines. Homology with the catalytic subunit of phosphatidylinositol 3-kinase (PI -3K) is drawn in aqua lines. Pink horizontal lines indicate the catalytic sites of PI-3K, DxxxxN and DFG. The very C-terminal gray lines indicate the FATC domain (FRAP, ATM, TRRAP C-terminal) (Bosotti et al., 2000).
Figure 6.3 LxxLL motifs present in AtTra1-2 and AtTra1-4.

The first conserved leucine residue in motifs is assigned +1. The conserved +1, +4 and +5 leucine residues are highlighted in blue and underlined. Hydrophobic residues at –1 (important in nuclear receptor binding) are black characters in gray background. Small or polar residues at –1 are white characters in black background. * denotes two motifs from the same sequence, so we counted as one motif.
The FATC domains (upper panel) from AtTra1-2 and AtTra1-4, and its homologs from yeast (Tra1), mammal (TRRAP), Drosophila (gi|62471663), C. elegans (gi|51011815) were aligned with the FATC domains (lower panel) from yeast TOR1 and TOR2 (the target of rapamycin), and its homologs from S. pombe (gi|19113067), M. musculus (gi|4826730), R. norvegicus (gi|561858), Drosophila (gi|17864562), and Arabidopsis (gi|22330143) using ClustalW. Residues in black background were highly conserved in the TOR proteins from different organisms. White lettering on black background indicates sequence identity, and white lettering on a grey background indicates sequence similarity. Note two critical Cys residues are completely conserved in each TOR protein (lower panel).
Figure 6.5 Molecular characterizations of *AtTra1* T-DNA insertion mutants.

(A) Schematic of T-DNA insertions in *AtTra1-2*. T-DNA insertions are in Exon 26 (D) and Intron 27 (G) respectively.

(B) Schematic of T-DNA insertions in *AtTra1-4*. T-DNA insertions are in Exon 7 (L), Exon 24 (K), and Exon 26 (N) respectively.

(C) RT-PCR analyses of *AtTra1-2D* and *-2G* mutants. PCRs were conducted using total RNA isolated from homozygous T-DNA insertion lines.

(D) RT-PCR analyses of *AtTra1-4K*, *-4L*, *-4N* mutants. PCRs were conducted using total RNA isolated from homozygous T-DNA insertion lines. Primer pairs used to determine expression are indicated above (forward primer) and below (reverse primer) in each schematic in A and B. RT-PCR analysis of *eIF4A* (primer pairs S140/S141) was used as internal control.
Figure 6.6 Expression analyses of *CBF1*, *ABF1* and *COR* genes in three *AtTra1-4* T-DNA insertion mutants.

The seedlings of wild type (WT) and three T-DNA insertion lines (AtTra1-4N, -4L, -4K) were cold treated at 4°C for various times. Ten ug total RNAs of pooled samples of more than 10 seedlings were loaded onto each lane. The same filter was hybridized, stripped, and re-hybridized subsequently with labeled probes for *CBF1*, *ABF1*, *COR47*, *COR6.6*, *COR78* and *COR15*. Ethidium bromide stained RNA is used as a loading control.
Figure 6.7 RT-PCR expression analyses of *AtTra1-2* and *AtTra1-4* in Arabidopsis tissues.

RT-PCR was conducted using AtTra1-2 gene specific primers S17 and S122; AtTra1-4, S211 and S118 (Table 6.1). Rt, root; R.L., rosette leaves; St, stem; C.L., cauline leaves; FB, flower buds; Slq, siliques. *eIF-4A* RT-PCR products serve as an internal control.
Figure 6.8 Representative genotypic analyses of AtTra1-2D x AtTra1-4K F₂ population.

T-DNA left border primer e916 and gene specific primers for AtTra1-2D or AtTra1-4K were combined for each reaction. Results shown here are representative of 160 F₂ individuals. D = AtTra1-2 WT, d = AtTra1-2D, K = AtTra1-4 WT, k = AtTra1-4K.
Figure 6.9 Phenotypes of T-DNA insertion mutant and wild type plant.

Arabidopsis wild type (WT) and T-DNA insertion mutant (d/d K/k) were sown in soil. Phenotypes were documented after four weeks growth in the greenhouse at 21~23°C with 14h light/10h dark.
Figure 6.10 Segregation analysis of embryo lethality associated with AtTra1 mutations.

(A) The numbers of normal and aborted embryos recovered from F3 individuals heterozygous for one AtTra1 and homozygous for the other. Siliques were opened and seeds were counted under a dissecting microscope.
(B) Genotyping of F4 progenies derived from selfed AtTra1-2+/--AtTra1-4-- F3 plants. Primers e936/e937 detect WT AtTra1-2; e916/e937 detect the AtTra1-2 T-DNA insertion; e920/e921 detect WT AtTra1-4; e916/e921 detect AtTra1-4K T-DNA insertion.
(C) Genotyping of F4 progenies derived from selfed AtTra1-2--AtTra1-4+/-- F3 plants.
(D) and (E) Illustration of primers used for genotyping. Primer e916 is within the T-DNA left border. E26 and E24 are exons 26 and 24 of AtTra1-2 and AtTra1-4, respectively.
Figure 6.11 Disruption of \textit{AtTra1-2} and \textit{AtTra1-4} results in embryo lethality.

Genetic crosses were made between the homozygous single gene mutant lines \textit{AtTra1-2D} and \textit{AtTra1-4K}. Opened siliques were photographed using a dissecting microscope. d/d K/k indicates homozygous \textit{AtTra1-2} mutant allele and heterozygous \textit{AtTra1-4} mutant allele; D/d k/k, heterozygous \textit{AtTra1-2} allele but homozygous \textit{AtTra1-4} allele. WT= D/D K/K. White and brown defective seeds (arrows) exist in one silique.
Figure 6.12 Alignment of *AtTra1-2* and *AtTra1-4* genomic regions with poplar cDNA clone PO 02035A02.

Poplar cDNA sequence was aligned with the putative promoter region and the first 420 coding region nucleotides of *AtTra1-2* and *AtTra1-4*, using ClustalW. The residue A in the predicted start codon ATG (highlighted in blue and underlined) is numbered +1. Minus 665 to −1 of *AtTra1-2* was fused to β-glucuronidase. Minus 724 to −10 of *AtTra1-4* was used fused to β-glucuronidase.
<table>
<thead>
<tr>
<th>P populous</th>
<th>P populous</th>
<th>P populous</th>
<th>P populous</th>
<th>P populous</th>
<th>P populous</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtTra1-2</td>
<td>AtTra1-2</td>
<td>AtTra1-2</td>
<td>AtTra1-2</td>
<td>AtTra1-2</td>
<td>AtTra1-2</td>
</tr>
<tr>
<td>+49</td>
<td>+49</td>
<td>+49</td>
<td>+49</td>
<td>+49</td>
<td>+49</td>
</tr>
<tr>
<td>+109</td>
<td>+109</td>
<td>+109</td>
<td>+109</td>
<td>+109</td>
<td>+109</td>
</tr>
<tr>
<td>+229</td>
<td>+229</td>
<td>+229</td>
<td>+229</td>
<td>+229</td>
<td>+229</td>
</tr>
<tr>
<td>+289</td>
<td>+289</td>
<td>+289</td>
<td>+289</td>
<td>+289</td>
<td>+289</td>
</tr>
<tr>
<td>+349</td>
<td>+349</td>
<td>+349</td>
<td>+349</td>
<td>+349</td>
<td>+349</td>
</tr>
<tr>
<td>+387</td>
<td>+387</td>
<td>+387</td>
<td>+387</td>
<td>+387</td>
<td>+387</td>
</tr>
<tr>
<td>+420</td>
<td>+420</td>
<td>+420</td>
<td>+420</td>
<td>+420</td>
<td>+420</td>
</tr>
</tbody>
</table>
Figure 6.13 β-glucuronidase (GUS) staining of Arabidopsis plants transformed with the AtTra1-2:GUS fusion.

(A) and (B) GUS activity in young leaves and young floral buds.
(C) Young bud but not developed buds showing GUS staining.
(D) Nodes of an auxiliary shoot in a mature plant.
(E) 20-day-old seedling.
YB, young bud; YL, young leaf; Nd, node; Rep, receptacle.
Figure 6.14 Schematic representation of two RNAi constructs.

(A) C-terminal end of Tra1/TRRAP has protein-protein interactions with many trans-activators. The PI3K used for RNAi construct was illustrated.

(B) Schematic representation of two RNAi constructs, controlled by either a constitutive CaMV35S RNA promoter (pZW146) or a cold-inducible \textit{COR15a} promoter (pZW169). Hairpin fragments were amplified from cDNA corresponding to the conserved PI3K of \textit{AtTra1-2} gene.
Figure 6.15 The phenotypes of RNAi lines.

(A) 35S RNAi seedlings (146) showed elongated hypocotyls (arrows) and abnormal cotyledons compared to control (either COR15a lines or lines transformed with vector only).

(B) RT-PCR analyses of AtTra1 gene expression. AtTra1-2 specific primers ES107/S17 amplify the PI3Klike domain. Primers S20/S21 amplify a region 2kb upstream of the PI3K domain of both genes (See figure 6.1B). Primers S140/S141 amplify eIF4A as an internal control. WT and RNAi line (169) were placed at 4 °C for 11 h before the extraction of total RNA, RNAi line (146) was not cold treated.

(C) Hypocotyl length of WT, COR15a RNAi line (169), and 35S RNAi lines (146-2 and 146-6). Each of twenty T2 seedlings was determined. The error bar is the standard deviation of the mean.

(D) 35S RNAi lines grown on selectable medium at 22 °C under 16h light/ 8h dark cycle. T3 seedlings expressing the RNAi construct (pZW146) were visually identified by the presence of an elongated hypocotyl. Both these individuals and the vector control were then transferred to new plates with selectable medium and grown alongside for an additional 16 d before being photographed.
Figure 6.16 Expression analyses of cold inducible genes in RNAi lines.

(A) RNA gel blot analysis of CaMV35S promoter (146) and COR15a promoter (169) driven Tra1 RNAi lines and vector control. Plants were placed at 4°C for the time points (in h) indicated above each lane. Total RNA was extracted from pooled seedlings with more than 10 individual transgenic seedlings.

(B) RNA gel blot analysis of COR15a promoter driven Tra1 RNAi lines. T1 seedlings were used in the cold temperature treatment.

(C) RNA gel blot analysis of CaMV35S promoter driven Tra1 RNAi lines and wild type control. T2 seedlings were used in this cold temperature treatment. Filters in A and B were hybridized, stripped, and re-hybridized in succession with probes for Tra1 PI3K region, COR47 and COR6.6. Filter in C was similarly probed with CBF1, ABF1, COR47, COR6.6, COR78, and COR15a. Ethidium bromide stained RNA was used as an internal control.
BIBLIOGRAPHY


Brown, C.E., Lechner, T., Howe, L., and Workman, J.L. (2000). The many HATs of

Brown, C.E., Howe, L., Sousa, K., Alley, S.C., Carrozza, M.J., Tan, S., and
Workman, J.L. (2001). Recruitment of HAT complexes by direct activator
interactions with the ATM-related Tra1 subunit. Science 292, 2333-2337.

activation domains stimulate initiation and elongation at different times and via
different residues. Embo J 17, 3146-3154.

and analysis of the Arabidopsis thaliana BSH gene, a member of the SNF5 gene

domains. Trends Genet. 6, 36-40.

DNA methylation, vernalization, and the initiation of flowering. Proc Natl Acad
Sci U S A 90, 287-291.

Plant Mol Biol 37, 425-435.

Callebaut, I., Labesse, G., Durand, P., Poupon, A., Canard, L., Chomilier, J.,
information through hydrophobic cluster analysis (HCA): current status and

effects of acidic activators on large-scale chromatin structure and transcription.
Mol Cell Biol 25, 958-968.


Nat Biotechnol 18, 602-603.

Ann Bot (Lond) 95, 901-915.


AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. Plant Cell 17, 3470-3488.


Kumar, R., Betney, R., Li, J., Thompson, E.B., and McEwan, I.J. (2004). Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. Biochemistry 43, 3008-3013.


Segal, D.J. (2002). The use of zinc finger peptides to study the role of specific factor binding sites in the chromatin environment. Methods 26, 76-83.


Steward, N., Kusano, T., and Sano, H. (2000). Expression of ZmMET1, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in actively proliferating cells, but also with altered DNA methylation status in cold-stressed quiescent cells. Nucleic Acids Res 28, 3250-3259.


Xue, G.P. (2002). An AP2 domain transcription factor HvCBF1 activates expression of cold-responsive genes in barley through interaction with a (G/a)(C/t)CGAC motif. Biochim Biophys Acta 1577, 63-72.


232


