Protein-binding sequences that may regulate organ-preferential transcription of zrp2, a Zea mays root-preferential gene

Therese Shamala Tirimanne
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/rtd
Part of the Botany Commons, and the Molecular Biology Commons

Recommended Citation
Tirimanne, Therese Shamala, "Protein-binding sequences that may regulate organ-preferential transcription of zrp2, a Zea mays root-preferential gene" (1995). Retrospective Theses and Dissertations. 11963.
http://lib.dr.iastate.edu/rtd/11963

This Dissertation is brought to you for free and open access by Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Protein-binding sequences that may regulate organ-preferential transcription of zrp2, a Zea mays root-preferential gene

by

Therese Shamala Tirimanne

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of DOCTOR OF PHILOSOPHY

Department: Botany Interdepartmental Major: Plant Physiology

Iowa state University Ames, Iowa

1995
# TABLE OF CONTENTS

**ABSTRACT** .......................................................................................................................... vi

**CHAPTER 1: INTRODUCTION** ................................................................................................. 1

Regulation of Gene Expression .......................................................................................... 1

Components Involved in Regulation of Transcription ....................................................... 2

- Cis-acting elements ......................................................................................................... 2
- Trans-acting transcription factors .................................................................................. 6
- DNA-binding domain of transcription factors ................................................................. 7

Regulation of Transcription ............................................................................................. 8

- Regulation of transcription by alterations in DNA-protein interactions .......... 8
- Regulation of transcription by alterations in protein-protein interactions .... 9
- Complexity of DNA-protein interactions ................................................................. 9
- Organ-preferential accumulation of mRNA ............................................................. 10

Methods of Analysis ........................................................................................................ 11

- Detection of DNA-protein interactions ................................................................ 11
- Analysis of transcriptional activity .......................................................................... 12

Goal of The Study ............................................................................................................ 13

**CHAPTER 2: INTERACTION BETWEEN THE OCS-ELEMENT OF THE OCTAPINE SYNTHASE GENE AND THE OCS-BINDING FACTOR IN MAIZE AS A MODEL SYSTEM FOR DNA-PROTEIN BINDING ANALYSIS** ................................................................. 15

Introduction ......................................................................................................................... 15

Materials and Methods ...................................................................................................... 17

Growth and harvest of seedlings ..................................................................................... 17
CHAPTER 3: ANALYSIS OF THE DNA-PROTEIN INTERACTIONS IN THE PROXIMAL REGION OF THE ZRP2 PROMOTER

Introduction .................................................................................................................. 26
Materials and Methods .................................................................................................... 29
Growth and harvest of seedlings ......................................................................................... 29
Isolation of nuclei and nuclear protein extract preparation .................................................. 29
Preparation of the probe and the competitor DNA ................................................................ 29
Sub-cloning of the first kbp of the zrp2 promoter ................................................................. 29
DNA fragments spanning the junctions ............................................................................... 31
Sub-division of the -472/NsiI to -180/SalI DNA fragment .................................................. 31
Deletion of the protein-binding region and cloning ............................................................... 31
Sub-division of the -379/NheI to -258/TaqI DNA fragment ............................................... 35
Cloning of the oligonucleotides spanning the -330 to -257 region ...................................... 35
Preparation of the probe and the specific competitor DNA ........... 37
Preparation of a non-specific competitor DNA ......................... 39
Purification of DNA ................................................................. 39
Gel mobility-shift assay ....................................................... 40
Results ......................................................................................... 41
DNA-protein interactions of the DNA fragments spanning the 1 kbp promoter ............................................................. 42
Specificity of the DNA-protein interactions ............................. 44
Sub-division of the -472/NsiI to -180/SalI DNA fragment to delimit the protein-binding region ....................................................... 50
Discussion ................................................................................. 60

CHAPTER 4: DETERMINATION OF THE NUCLEOTIDE SEQUENCE OF THE UPSTREAM REGION OF THE ZRP2 PROMOTER .......................................................... 72
Introduction .................................................................................. 72
Materials and Methods ................................................................ 74
Sub-cloning and sequencing of the -976/EcoRV to -1945/HindIII region ............................................................. 74
Sub-cloning and sequencing of the -1945/HindIII to -2573/Apal region ............................................................. 75
Results ......................................................................................... 77
Discussion ................................................................................. 82

CHAPTER 5: EFFECT OF THE PROTEIN-BINDING SEQUENCES ON TRANSCRIPTION OF THE ZRP2 PROMOTER .......................................................... 84
Introduction .................................................................................. 84
Gene-transfer systems ................................................................ 85
Transformation by microprojectile bombardment ....................... 85
ABSTRACT

In this study, DNA-protein interactions of a one kilobase region proximal to the transcription start site of the zrp2 (Zea mays root-preferential 2) gene promoter was analyzed. The intent of this research was to identify regions of the promoter and DNA-protein interactions, that are responsible for root-preferential transcription. In vitro run-on transcription assays had revealed dramatically higher levels of transcription in 3-d-old roots than in 5-d-old leaves (Held, 1993). Interactions between regions of the zrp2 promoter and nuclear protein extracts from 3-d-old roots and 5-d-old leaves of maize were analyzed by gel mobility-shift assays. Two DNA-protein complexes with each protein extract were observed with the -472 to -180 region of the promoter. This region showed 72% nucleotide identity with a 159 bp portion of the first intron from an α-tubulin gene that is expressed preferentially in the roots of maize (Montoliu et al., 1989). A DNA-protein complex with high electrophoretic mobility (Z2BP1) was observed with both root and leaf nuclear protein extracts. Complexes Z2BP2 and Z2BP3 were preferentially formed with root and leaf nuclear protein extracts, respectively. Subsequently, the protein-binding region was delimited to a 73 bp (-330 to -258) region, which contained several repeats of two sequence motifs consisting only of As and Ts. The effect of the protein-binding sequence on transcription of the 1 kbp and the 4.7 kbp promoter fragments (fused to the GUS reporter gene) were analyzed by particle bombardment-mediated transient assays in roots and leaves of maize. The transcriptional activities of promoters, with the protein-binding region deleted were compared with their respective intact promoters. In roots, the deletion resulted in 5-fold-less transcriptional activity with the 1 kbp promoter fragment, while the 4.7 kbp promoter showed no significant decrease. In leaves no significant decrease was observed with either 1 kbp or 4.7 kbp promoters. We conclude that this protein-binding region is important in enhancing the level of transcriptional activity in roots.
CHAPTER 1: INTRODUCTION

Regulation of Gene Expression

Plant development and function is critically dependent on the precise control of gene expression. In any one cell at a particular time, only a subset of the total genetic information is expressed. The available evidence shows that there are unique and overlapping sets of mRNAs among different organs of a plant (Kamalay and Goldberg, 1980, 1984). Consequently, the presence of a gene product may be limited to a certain cell type or organ at specific stages of development or in response to a distinct stimuli.

Expression of a gene is a multistep process, including messenger RNA synthesis in the nucleus, transport of newly synthesized mRNA to the cytoplasm, synthesis of protein in the cytoplasm and transport to other cellular locations, resulting in the production of a functional product in the proper location and time. Regulation can be applied at numerous points in this pathway. Initiation of mRNA synthesis is a primary control point in the regulation of differential gene expression. Cells respond to intra- and extra-cellular cues by turning the transcription of certain genes on or off and by modulating the amount of transcription. Although the mechanisms and biochemical pathways by which cells integrate physiological cues to yield appropriate transcriptional changes are still unknown, it is clear that the frequency of initiation of mRNA synthesis depends ultimately on protein factors that interact with discrete specific elements in promoters of genes.
Components Involved in Regulation of Transcription

Cis-acting elements

Eukaryotic organisms contain three distinct forms of nuclear DNA-dependent RNA polymerases, pol I, pol II and pol III, each responsible for the transcription of a particular class of genes. Although the three polymerases are structurally closely related, distinct mechanisms appear to be involved in transcription initiation by each enzyme (Sawadogo and Sentenac, 1990). Regulatory elements for RNA polymerase II transcription can be scattered both upstream and downstream of the transcription start site of a gene (Johnson and McKnight, 1989). Systematic mutational analyses have revealed that each gene has a particular combination of positive and negative regulatory cis-elements that are uniquely arranged in type, number and spatial distribution. These elements are binding sites for sequence-specific transcription factors that modulate transcription of that gene. Typically, cis-elements are arrayed within several hundred base pairs from the transcription initiation site. The control region (in the immediate vicinity of the transcription start site) which is capable of driving transcription is called the promoter. Promoters are composed of multiple cis-elements or modules, each contributing to a complex pattern of transcriptional regulation (Dynan, 1989). Regions that regulate a promoter from a distance in orientation independent manner are called enhancers (Ptashne, 1986).

Cis-acting elements can be broadly categorized into two classes: sequences that are binding sites for RNA polymerase complex (or for factors associated with RNA polymerase) that are essential for basic transcription (general factors), and sequences that are binding sites for protein factors that may interact with the RNA polymerase complex to regulate the level of transcription.
As the basic transcriptional apparatus is common to all pol II-transcribed genes, in principle sequence elements involved in the binding of RNA polymerase and general transcription factors should be conserved in all promoters. Sequence comparisons between different promoters show several short consensus sequences that are conserved. The TATA box, CAAT box and GC box are the most important and best characterized of these promoter-elements.

The TATA box is situated approximately 30 bp upstream of the transcription start site and has the consensus of TATAAT flanked by CG-rich sequences in plants. This element is recognized by a sequence-specific TATA-binding protein (TBP), which is a component of a large complex of proteins (TAFs) forming the basal transcription factor, TFIID. This factor functions in organizing the assembly of the transcription complex, including transcription factors TFIIA, TFIIB, TFIIE, TFIIF and RNA polymerase II on to the promoter (reviewed by Sawadogo and Sentenac, 1990). Mutations in the TATA box interfere with the proper transcription initiation, and promoters that lack TATA box usually lack unique start points.

The CAAT box is preserved only in some promoters and is usually located around 80 bp upstream of the transcription start site. Plant genes frequently do not contain a CAAT box. Nevertheless in cereal promoters, a CATC sequence located at -90 bp may serve as the substitute for the CAAT box (Kries et al., 1986). The GC box contains the sequence GGGCGG, and often is present in multiple copies. Both the CAAT box and the GC box can function in either orientation, and are believed to play a key role in bringing the RNA polymerase into the vicinity of the transcription start site, thus influencing the frequency of initiation.
Cis-acting elements that are binding sites for proteins interacting with the RNA polymerase complex can be of several types. Through interaction with specific transcription factors, they can confer location-, time- and developmental stage-specific/preferential transcription or regulate signal-dependent transcription. Most genes transcribed in an organism are regulated in a cell, tissue, organ, developmental stage and environmental cue-specific manner. Cells respond to environmental or developmental stimuli by switching transcription on, off or by altering the level of transcription of respective genes. Cis-elements through interaction with transcription factors act as enhancers or repressors of transcription. Although there are a plethora of examples of enhancer elements, not many repressor elements have been identified in plants (Bruce et al., 1991; Haaren and Van Houck, 1991).

Cis-elements are responsible for conferring cell-, tissue- or organ-specific transcription in plants (Benfy et al., 1989, 1990a; Leva et al., 1992). Promoter deletion analyses have revealed that promoters consist of an array of cis-elements, some capable of organ- or tissue-specific transcription. The Cauliflower Mosaic Virus (CaMV) 35S promoter, although functioning as a constitutive promoter in total, directs transcription of the GUS reporter gene specifically in roots, when deleted up to -90. The region from -343 to -90 directs transcription in shoots (Benfy et al., 1989). In the bean phenylalanine ammonia-lyase gene 2 (PAL2) promoter, cis-elements located between -289 to -74 from the transcription start site are essential for transcription of GUS in the xylem and -480 to -289 region for expression in the phloem. (Levya et al., 1992). Groups of genes that are regulated by a common external factor generally share common cis-acting regulatory elements that are responsible for the signal-dependent transcriptional regulation. A large number of cis-elements (response elements) have been implicated in specialized types of
signal-dependent transcriptional regulation. Light response elements (LRE), heat shock response elements (HSE), growth hormone response elements, such as ABA response element and auxin response elements, are some examples of such cis-regulatory elements in plant genes.

Typically, cis-elements are located within several hundred base pairs from the transcription initiation site, but some elements can exert control over distances as great as 30 kbp (Mitchell and Tjian, 1989). Protein factors bound to cis-elements located in the proximal region, interact with the RNA polymerase complex directly, or indirectly, through protein bridges (Herendeen et al., 1990; Müller et al., 1989). Even if a cis-element is several hundred base pairs away, physically it may be located close to the transcription initiation site due to the topology of the DNA (Laybourn and Kadonaga, 1992). Cis-elements located several kilo bases away from the transcription initiation site (long-range regulatory elements) are known to participate in transcription initiation by tracking of a protein along DNA, or by association of two proteins bound at separate sites to form a DNA loop in between or by interactions that are affected by the topology of DNA (Wang and Giaever, 1988; Müller et al., 1989).

The importance of a particular cis-element can vary greatly in different cell types or under different physiological conditions as the DNA-binding factor or factors that recognize the element may vary in abundance or in the ability to function, in different tissues, or under different conditions. Overlapping binding sites for different factors can result in multiple factors competing for binding sites, allowing regulation according to the abundance of a factor, or the binding affinity of a factor to the cis-element. Duplication of a regulatory sequence can increase the strength of an enhancer (Kay et al., 1987). AT- and GC-rich stretches of DNA, and direct or inverted repeat sequences, have been shown to
have a high probability of interacting with protein factors regulating transcription. In all cases, a specific protein factor interacts with the specific cis-element to cause regulation of transcription. As the same cis-elements are present in the DNA of every cell, for differential regulation of transcription, trans-acting proteins should only be available, or active under the conditions, where the gene is to be transcribed.

**Trans-acting transcription factors**

DNA-binding proteins play a central role in biology. Among other activities, they are responsible for replicating the genome, for transcribing active genes, and for repairing damaged DNA. One of the largest and most diverse classes of the DNA-binding proteins are the transcription factors, that regulate transcription by binding to a specific DNA sequence.

Trans-acting protein factors that interact with cis-elements of promoters can be categorized into two classes according to their function in transcription. Protein factors that are necessary for basic transcription are called general transcription factors. Transcription factors TFIIA, TFIB, TFIID, TFIIE, TFIIF and TBP associated factors (TAFs) are examples of general transcription factors that are present in all cells of an organism. The other category of DNA-binding proteins are present, or active, only in cells under a specific condition or in a specific location or time, and are strictly regulated. These proteins have the ability to recognize their specific target DNA sequences and to interact with the RNA polymerase complex directly or indirectly, through protein-protein interactions with other transcription factors. For a repressor protein, binding to a DNA sequence itself may be sufficient to exercise its function. Protein-DNA interaction depends on the secondary and the tertiary configurations of the protein, which enable it to be
compatible with the DNA surface it binds to. DNA-binding activity is localized in the 
DNA-binding domain of the protein, and transcriptional-activating functions are localized 
in the transcriptional activation domain of the protein. In several mammalian factors, these 
domains consist of separate, 30- to 100-amino acid sub-regions (Mitchell and Tjian, 1989).

**DNA-binding domain of transcription factors**

In general, proteins bind to the B form of DNA (for review see Glover, 1989; Pabo and 
Sauer, 1992). In many of the systems structural differences have been observed between 
the B form of DNA and the DNA in the complex, and between the free protein and the 
bound state of the protein. Binding may induce a significant distortion such as bending or 
kinking in the DNA (Pabo and Sauer, 1992), or folding in the protein (Travers, 1989). The 
role of this conformational change may be to cause an 'induced fit' between the specific 
DNA and the protein (Spolar and Record, 1994). Proteins interact primarily with the major 
and/or the minor groove of the DNA through hydrogen bonding, ionic interactions and 
hydrophobic interactions. Most commonly, hydrogen bonds form between protein side 
chains or polypeptide backbone and DNA bases, mediated by water molecules. Strength or 
energy of the interaction increase with the number of contact sites, and, therefore, when 
protein dimers or trimers bind palindromic sequences of DNA, the interaction becomes 
more stable.

Structural studies and sequence comparisons of many transcription factors show 
common types of motifs that are responsible for binding to DNA (Pabo and Sauer, 1992). 
Based upon these sequence motifs, DNA-binding proteins have been classified into several 
classes. Large, well established classes include helix-turn-helix proteins, the 
homeodomains, zinc finger proteins, the steroid receptors, leucine zipper proteins and the
helix-loop-helix proteins (Glover, 1989). In addition there are other classes that are not yet well-characterized. The bZIP proteins and MADS proteins, which include several plant transcription factors (Kuhlemeier, 1992; Lohmer et al., 1991; Oeda et al., 1991; Coen and Meyerowitz, 1991; Pnueli et al., 1991), fall into this category.

**Regulation of Transcription**

**Regulation of transcription by alterations in DNA-protein interactions**

One means of regulating transcription is through modulation of the binding ability between the cis-element and specific transcription factors. As the abundance of cis-elements are more or less constant, the accessibility of the cis-element, affinity of the interacting site, and the availability of active transcription factors are more commonly subjected to regulation.

A change in the availability or the abundance of a specific active transcription factor can be caused by altering the synthesis or degradation of the factor, or by modification of the already existing protein factor (Hunter and Karin, 1992; Datta and Cashmore, 1989). Alterations in the conformation of the DNA, such as bending, kinking, or unwinding (Perez-Martin and Espinosa, 1993; Gartenberg and Crothers, 1988; Pabo and Sauer, 1992), can influence the accessibility to the binding site, or the binding affinity. This may facilitate or hinder the binding of the sequence-specific transcription factors to the cis-elements during transcription.
Regulation of transcription by alterations in protein-protein interactions

Another means of regulating transcription is through activation or modification of the transcription factor. This can be caused by phosphorylation or specific protein-protein interactions of the transcription factor with other associated proteins and ultimately with the RNA polymerase II complex. Alterations in the accessibility of the activation domain (Jayaraman et al., 1994), or alterations in the affinity of the protein-protein interactions play a major role in changing the activation potential of the transcription factor in transcription. However, although there are many examples of protein-protein interactions promoting or preventing DNA-binding by transcription factors, much less is known of mechanisms underlying this mechanism of transcription activation. This lack of information arises both because of the complexity of the transcription activation process and from the difficulties involved in reconstitution of transcription regulation using purified components with plant or mammalian systems.

Complexity of DNA-protein interactions

Binding of transcription factors to their specific cis-elements is complex. A single transcription factor may bind to one or more copies of a single cis-element, or multiple cis-elements. Conversely more than one factor, usually belonging to the same family of transcription factors, may bind to a single cis-element or to multiple cis-elements.

The cis-element activation sequence-I (as-1, Lam et al., 1989), situated in the promoter region (-83 to -63) of the cauliflower mosaic virus 35S promoter, can bind to the nuclear factor ASF-1 and also to the cloned transcription factor TGA1a (Katagiri et al., 1989). Two other cis-elements, octapine synthase (ocs) and nos-1, from the promoter regions of Agrobacterium opine synthase genes can also interact with the same factors.
(Fromm et al., 1989; Lam et al., 1990). If multiple factors can bind to a specific cis-element, and if these factors are present in the same cell, competition between factors for binding may occur. Moreover, if a single factor can bind to multiple cis-elements, competition between cis-elements for binding to the factor may occur (Neuhaus et al., 1994). In situations like these, mechanisms must exist to maintain the regulatory specificity required for the precise and coordinated regulation of transcription. The presence of multiple copies of a single cis-element, or the relative abundance of a specific factor, or high binding affinity between the cis-element and the factor enhance the chance of finding its cognate factor. Co-operative binding of factor(s) to cis-elements facilitates the binding of specific factors even if other combinations are possible.

**Organ-preferential accumulation of mRNA**

Organ-preferential or specific accumulation of mRNA can be due to differential transcriptional and/or post-transcriptional (RNA processing, transport, degradation) regulation. Nuclear run-on experiments using nuclear extracts of different organs can be used to determine whether differential accumulation of transcripts is due to regulation at the level of transcription. Different approaches have been used for analysis of the promoter region for cis-elements associated with organ-preferential/specific transcription. Promoter deletion/reporter-gene fusion constructs in transient or in stable-transformation-expression studies can indicate the region of the promoter containing cis-elements important in driving organ-preferential transcription. Analysis of promoter regions for DNA-protein interactions, using nuclear extracts of different organs can reveal location of cis-elements on DNA, and the pattern of the interaction. Generally a correlation between the presence or the abundance of a DNA-protein complex, with the transcriptional activity of its target
gene may provide circumstantial evidence for its involvement in the regulation of transcription (Forde, 1994; Cushman and Bohnert, 1992). A different pattern of protein-DNA interactions in different organs where transcription is differentially regulated may indicate specific cis-elements involved in organ-preferential/specific transcription (Manzara et al., 1991). Several cis-elements responsible for organ-specific/preferential transcription have been identified by the presence of such correlation between the pattern of DNA-protein interactions and differential transcription of the mRNA between two organs (Manzara et al., 1991; de Pater et al., 1993; Meller et al., 1993). If low transcription is due to low abundance of the transcription factor, injection of the cloned factor into the cells of the organ where transcription is limited may show increased transcription (Neuhaus et al., 1994), confirming that a particular factor interacting with the specific cis-element is responsible for the organ-specific transcription.

**Methods of Analysis**

**Detection of DNA-protein interactions**

Protein factors binding to particular cis-elements have been detected by gel mobility-shift assays and DNase I footprint assays. In a gel mobility-shift assay, a labeled DNA fragment containing the specific cis-element is incubated with a nuclear protein extract and then analyzed on a non-denaturing gel. A protein-DNA complex migrates slower than free DNA. The specificity of the interaction can be determined by adding excess unlabeled DNA to the binding reaction. DNA sequences related to the cis-element compete for binding and unrelated DNA sequences do not. Most DNA-protein interactions characterized in plants so far have been detected using gel mobility-shift assays (Allen et

It can be concluded from gel mobility-shift assays that a particular fragment of DNA contains cis-element(s) that interact with protein factor(s). All factors that bind DNA are not involved in regulation of transcription. Therefore, in order to determine whether a DNA-protein complex is involved in regulation of transcription, the transcriptional activity of the promoter with and without that particular cis-element could be investigated.

Analysis of transcriptional activity

Transcriptional activity of a promoter can be investigated by fusion to a reporter gene and by introducing the heterologous gene construct into the plant using a transient or stable transformation system. The transcriptional activity is measured by the activity of the product of the reporter gene. For this, it must be assumed that the activity of the gene product is directly reflecting the transcriptional activity and hence is a measure of the transcriptional activity. For this assumption to be correct, post-transcriptional regulation should not occur. As a reporter gene, a β-glucuronidase (GUS) coding sequence is widely used in plants (Jefferson et al., 1987). This gene is originally from bacteria, and the enzyme product is absent in plants. Moreover, the active enzyme can cleave a wide range of β-glucuronides, many of which are commercially available as spectrophotometric, fluorometric, and histochemical substrates.

Investigation of the effect of a cis-element on transcription can be done by gain-of-function or loss-of-function studies. Several copies of the putative cis-element (enhancer element, as an example) are fused to the 5' end of a minimal promoter/reporter-gene fusion construct and compared with the minimal promoter activity (for increased transcriptional
activity) in gain-of-function studies. On the other hand, in loss-of-function studies the putative cis-element is deleted, and the transcriptional activity is compared with the intact promoter. Changes in the transcriptional activity would disclose whether the cis-element is involved in transcriptional regulation and whether its effect is positive or negative. If the exact protein-binding site is known, alterations in the bases of the binding site could be done with point mutations instead of deleting the whole cis-element. This is advantageous as promoters with different point mutations would show different levels of transcriptional activity and, therefore, can be compared with the effect of mutations in DNA-protein-binding activity in gel mobility-shift assays. A correlation between the binding activity and the transcriptional activity provides circumstantial evidence for the involvement of the DNA-protein complex in regulation of transcription.

**Goal of The Study**

The primary goal of this study was to identify regions of the zrp2 (*Zea mays* root-preferential gene 2) promoter that may be responsible for (or associated with) the root-preferential transcription of this gene. Initially, gel mobility-shift assays were performed with DNA fragments spanning the first kbp of the zrp2 promoter, proximal to the transcription start site, to investigate whether there are any differences in DNA-protein interactions between leaves and roots, that may be responsible for differential transcription between the two organs. Once a region of the promoter that showed differential binding was discovered, attempts were made to delimit the protein-binding site to a smaller region. To determine whether this region was involved in transcriptional regulation, the transcriptional activities of the intact 1 kbp promoter and the promoter with protein-
binding site deleted were compared, using particle bombardment-mediated transient assay system.

In general, promoters capable of directing transcription in monocotyledonous plants are very useful in genetic engineering of agronomically important cereals, as most plant promoters from dicotyledonous plants do not function as well in monocots (Kyozuka et al., 1993). The study of the cell or organ-specific transcription of plant genes in vegetative organs has numerous applications to agricultural biotechnology. For example, plant promoters that direct vascular specific gene expression can be used to express products that protect against phloem or xylem-borne pathogens. A trichome-specific promoter could be used to express novel anti-pest products in trichomes. Likewise a root-specific or preferential promoter could be used to express anti-pest products that protect against root-borne pathogens.
CHAPTER 2: INTERACTION BETWEEN THE OCS-ELEMENT OF THE OCTAPINE SYNTHASE GENE AND THE OCS-BINDING FACTOR IN MAIZE AS A MODEL SYSTEM FOR DNA-PROTEIN BINDING ANALYSIS

Introduction

Interaction between an upstream sequence of the octapine synthase gene (ocs-element) and a protein factor from nuclear protein extract of a cell suspension culture in maize was chosen as a model system to investigate, prior to the analysis of DNA-protein interactions of a promoter of a root-preferential gene zrp2, in maize. This system was chosen mainly due to the presence of the mRNA encoding the binding factor in young roots and leaves of maize (Singh et al., 1990), which suggested the possibility of the presence of the binding factor in root and leaf nuclear protein extracts. Analysis of the binding activity of crude nuclear protein extract preparations of roots and leaves, with the ocs-element, showed the capability of protein factors in the nuclear protein extract preparations to bind to the specific DNA sequence, allowing the use of the same nuclear protein extracts from roots and leaves for binding assays with the zrp2 promoter.

The ocs-element is a transcriptional enhancer, that has been shown to be a functional component of a number of genes that are active in plants (Bouchez et al., 1989). It was first characterized as a 16 bp palindromic sequence (ACGTAAGCGCTTACGT), situated in the upstream sequence of the octapine synthase gene (ocs) in Agrobacterium tumefaciens T-DNA (Ellis et al., 1987). The ocs-elements have also been found in other T-DNA encoded opine synthase genes and in the promoters of three plant DNA viruses,
including Cauliflower Mosaic Virus (CaMV). Comparison of the sequences of the ocs-element in these genes show a 20 bp consensus sequence, which includes the 16 bp palindrome (Bouchez., 1989). During infection, the ocs gene is transferred and integrated into the genome of the plant cells, with the T-DNA of Agrobacterium. The ocs gene is not expressed in Agrobacterium, but is expressed in the infected plants (Otten et al., 1981). Although the infectivity of Agrobacterium is generally limited to dicots, the transcriptional enhancer of the ocs promoter functions in both dicots and monocots and does not require any factors supplied by other genes of the Agrobacterium (Ellis et al., 1987). The 16 bp ocs-element, when inserted upstream of a reporter gene containing a truncated minimal promoter, enhances expression of the reporter gene in transient assays with protoplasts of tobacco and maize (Ellis et al., 1987).

A protein that binds to the ocs-element has been identified in nuclear protein extracts from maize cell suspension cultures (Tokuhisa et al., 1990). Gel mobility-shift assays and DNase footprint analysis have shown formation of two DNA-protein complexes. Interaction of the protein with both halves of the palindrome results in a low mobility complex, while binding of the protein to one half of the palindrome results in a high mobility complex (Singh et al., 1989; Tokuhisa et al., 1990). Binding of the protein to both halves of the ocs-element is needed for transcriptional activation in vivo (Bouchez et al., 1989; Singh et al., 1989; Tokuhisa et al., 1990).

In this analysis crude nuclear protein extracts from young roots and leaves of maize were used, instead of cell suspension cultures, in gel mobility-shift assays. Binding patterns of the nuclear factors in crude extracts were similar to those observed with nuclear protein extracts of protoplast suspension cultures. These observations lead to the
conclusion that the crude nuclear protein extracts used contain factors capable of binding specifically to their cognate cis-element, under the conditions used in these experiments.

**Materials and Methods**

**Growth and harvest of seedlings**

Maize (*Zea mays* L. cv NKH31) seeds were surface sterilized in a 50% commercial bleach solution for ten minutes, rinsed three times with water and imbibed for one hour. Approximately 50 seeds were placed on germination paper, which was then rolled and placed vertically in a polyethylene container containing 5 to 8 cm of water. Seedlings were grown at 30°C under a 12 h light/12 h dark cycle in a growth chamber.

Roots from three-day-old seedlings and leaves from five-day-old seedlings were harvested for nuclear protein isolation. Seedlings were placed on ice cold glass plates and roots were excised with a razor blade. Primary leaves were harvested by gradually pulling the leaf from the sheath. All plant materials were immediately chilled on ice after harvest.

**Isolation of nuclei**

Isolation of nuclei was carried out as in the method of Held (1993). Five grams of fresh roots or leaves were used per isolation of nuclei. All reagents were pre-cooled to 4°C and all steps were carried out at 4°C. Roots or leaves were immersed in ether for four or three minutes respectively. The ether was drained off and 5 ml of extraction buffer containing 2.5% ficoll 400, 4.0% dextran T40, 250 mM sucrose, 25 mM Tris-HCl, pH 7.8, 10 mM MgCl₂ and 4 mM phenylmethanesulfonyl fluoride (PMSF) was added to the mortar. Plant materials were thoroughly ground and the homogenate was filtered through
two layers of nylon meshes of 60 mm and 140 mm. The mortar and the meshes were rinsed with another 10 ml of extraction buffer. To the filtrate 500 ml of 20% Triton-X-100 was added and mixed, followed by centrifugation for 3 to 4 min at 3000 rpm in an HB-4 rotor. The supernatant was decanted, and the pellet was resuspended in 10 ml of wash buffer (extraction buffer + 0.1% Triton-X-100 vol/vol) and centrifuged again at 2000 rpm for 2 to 3 minutes. The pellet was resuspended in 300 ml of nuclear suspension buffer containing 15 mM Hepes pH 7.8, 100 mM KCl, 5 mM MgCl₂, and 5 mM 2-mercaptoethanol. Small aliquots were stained with methyl green and viewed under a light microscope.

**Preparation of crude nuclear protein extract**

Nuclear protein extracts were prepared by modification of the method described by Parker and Topol (1984) and Tokuhisa et al., (1990). Approximately equal numbers of nuclei from nuclear preparations of leaves and roots were used as determined by the number of nuclei per viewing area, under 500X magnification. Final volume of the nuclei preparations was adjusted to 8 ml with the nuclear suspension buffer and was transferred to polycarbonate ultracentrifuge tubes. Nuclei were lysed by adding 1 ml of 3.3 M (NH₄)₂SO₄/0.05 M Tris, pH 7.8, to adjust the ammonium sulfate concentration to 380 mM, and by gently rocking for 30 min at 4°C. The lysate was clarified by centrifugation at 55,000 rpm for 60 min in a Beckman 70.1 Ti fixed-angle rotor. The clear supernatant was transferred to another polycarbonate ultracentrifuge tube, and the proteins were precipitated by adding 0.25 g ammonium sulfate per ml of the supernatant. After mixing 30 min at 4°C, the supernatant was centrifuged for 15 min at 55,000 rpm in Beckman 70.1 Ti rotor. The supernatant was discarded, and the pellet was resuspended in a 1 ml of nuclear binding
buffer containing 10% glycerol, 25 mM Hepes pH 7.8, 50 mM KCl, 1 mM EDTA, and 50 mM 2-mercaptoethanol. The protein preparation was subsequently dialysed against two changes of 1.5 L nuclear binding buffer for 4 h. The dialyzed protein was subjected to centrifugation at 10,000 rpm for 15 min. Small aliquots of the extract were frozen in liquid nitrogen and stored at -80°C.

**Determination of the concentration of protein**

The protein concentration of the crude nuclear protein extract was determined according to the method of Bradford (1976), using bovine serum albumin (BSA) as a standard. Nuclear protein concentrations ranged from 2 to 3.5 mg/ml.

**Preparation of the probe and the competitor DNA**

**Cloning of the oligomer containing the ocs-element.** Manipulation of DNA was done essentially as described by Sambrook et al., (1989). Two single-stranded DNA oligomers (36 nucleotides in length and complimentary to each other), containing the ocs-element of the octapine synthase gene promoter were synthesized at the ISU Nucleic Acids Facility. Restriction sites EcoRI and HindIII were included at the ends, hence when the two strands were annealed, it would leave intact overhangs of the restriction sites to facilitate direct cloning into pBluescript. A NcoI site was included next to the ocs-element to use as a unique restriction site to screen for transformants (Fig. 2.1). Ten micrograms each of the oligomers were annealed by heating at 90°C for 15 min and cooling slowly to room temperature. The resultant double-stranded oligomer was subsequently cloned directly in to HindIII and EcoRI-digested pBluescript KS to construct pOCS(36), and the transformants were selected for the presence of NcoI site.
Figure 2.1 Sequence of the annealed oligomers containing the ocs-element. The sequence of the ocs-element is underlined and restriction sites are indicated.

Preparation of the probe and the specific competitor DNA. The plasmid pOCS(36) in pBluescript was double digested with BamHI and XhoI to obtain a 75 bp DNA fragment containing the ocs-element to use both as the probe and as the specific competitor DNA. The digestion reaction was end-labeled by the Klenow fragment of the DNA polymerase I, using 32P-labeled dCTP to produce the labeled probe. Total incorporation of radioactivity was determined by liquid scintillation spectrometry.
Preparation of a non-specific competitor DNA

A non-specific competitor DNA of 74 bp was obtained by the restriction digestion of pBluescript KS with KpnI and XbaI. This fragment included most of the pBluescript multiple-cloning site.

Purification of DNA

Probe, specific competitor DNA, and non-specific competitor DNA fragments were gel-purified by electrophoresis in 12% non-denaturing polyacrylamide gels (29:1 acrylamide to bis-acrylamide), with TBE (Tris-borate-EDTA) buffer at 80 volts for one hour. Fragments of the gel containing DNA were excised using an autoradiogram of the gel as the template. Polyacrylamide gel pieces containing the probe, specific competitor DNA, or non-specific competitor DNA were placed in 50-100 ml of nuclear binding buffer overnight, at 37°C, to facilitate diffusion of the DNA into the buffer.

Gel mobility-shift assay

Gel mobility-shift assays were performed as in Singh et al. (1989). Binding reactions contained $10^4$ cpm of end-labeled probe, 1.0 µg of poly(dI-dC).poly(dI-dC), 4 µg of nuclear protein extract in nuclear binding buffer. Total volume of a reaction mixture was 20 ml. Reactions were begun by the addition of labeled probe to the reaction mixture containing poly(dI-dC).poly(dI-dC) and nuclear protein extract. After mixing, reactions were incubated for 30 min at 30°C.

In competition assays, non-labeled probe was added to the reaction mixture and incubated at 30°C for 20 min, prior to the addition of labeled probe.

After incubation, reaction mixtures were electrophoresed in 4 to 7% non-denaturing polyacrylamide gels (80:1 acrylamide to bis-acrylamide) according to the size of the probe.
Gels were run at 70 volts for 1-1½ hours (10V/cm) at room temperature, using 200 mM Tris, 220 mM Borate and 5 mM Na₂EDTA (pH 8.3) as the running buffer. Following electrophoresis, gels were transferred onto 3M paper and dried at 80°C for 1 hr. Autoradiography was performed at -80°C with a single intensifying screen.

**Results**

Interactions between the *ocs*-element and protein factors in crude nuclear protein extracts of 3-d-old roots and 5-d-old leaves of maize were analysed by gel-shift assays. Binding reactions contained duplex poly(dI-dC).poly(dI-dC) to decrease non-specific interactions between DNA and proteins. Two DNA-protein complexes of slightly different mobility were observed with both root and leaf nuclear protein extracts (Fig. 2.2, lanes 2 and 3). The mobility of the complexes formed with both root and leaf nuclear protein extracts were similar.

To determine whether these DNA-protein interactions were specific, gel-shift assays, were performed with the *ocs*-element in the presence of specific and non-specific competitor DNA. Figure 2.3 shows results of the competition assays with leaf nuclear protein extracts. A non-labeled 75 bp DNA fragment containing the *ocs*-element was used as the specific competitor. As non-specific DNA, a non-labeled 74 bp DNA fragment from pBluescript multiple-cloning site (KpnI - XbaI) was used. Molar ratios of 1:10 and 1:20 probe to specific competitor DNA (lanes 3 and 4), and molar ratios of 1:20 probe to non-specific competitor DNA (lane 5), were used to compete with the probe in competition assays. Lane 3 exhibited a drastic decrease in the formation of complexes in the presence of increasing specific competitor amount in the reaction mixture. DNA-protein complexes
Figure 2.2 Gel-shift analysis of labeled ocs-element with nuclear protein extracts from leaves and roots. The end-labeled 75 bp DNA fragment containing the ocs-element (10,000 cpm), was incubated with (+) or without (-) 9 µg of nuclear protein extracts from 5-d-old leaves (+L) or 4 µg of nuclear protein extract from 3-d-old roots (+R), in the presence of 1 µg of double stranded poly(dI-dC).poly(dI-dC). The two DNA-protein complexes formed are indicated.
Figure 2.3  Competition assay using the DNA fragment containing the *ocs*-element with leaf nuclear protein extract. Gel-shift assays were performed with either free probe alone (lane 1), with no competitor DNA (lane 2), with 10x and 20x molar excess of non-labeled specific competitor DNA (lanes 3 and 4), and with 20x molar excess of non-specific competitor DNA (lane 5). After electrophoresis, the dried gel was subjected to autoradiography for 17 h at -80°C.
were fully competed away by a 1:20 molar ratio of the specific competitor DNA (lane 4). However, a 1:20 molar ratio of non-specific competitor DNA was not able to compete with the probe for binding with protein (lane 5). Failure to abolish or alter binding activity in the presence of molar excess quantities of non-specific competitor DNA shows that the interactions between the *ocs*-element and the protein are specific.

**Discussion**

Crude nuclear protein extracts from 3-d-old roots and 5-d-old leaves of maize contained a binding activity that shifted the mobility of the $^{32}$P labeled DNA probe containing the *ocs*-element. These gel mobility-shifts were similar to the data shown by Tokuhisa et al. (1990), with the nuclear fraction of maize cell suspension cultures. Similarly, an excess of non-labeled DNA fragments containing the *ocs*-element in binding reactions were able to compete with $^{32}$P labeled probe for binding (fig. 2.3, lanes 3 and 4). An excess of non-labeled, non-specific DNA fragments were not able to compete with $^{32}$P labeled probe (fig. 2.2, lane 5), indicating the specificity of the *ocs*-element-protein interaction.

These results demonstrate that the crude nuclear protein preparations of roots and leaves contain protein factors capable of binding specifically to the *ocs*-element under conditions the experiments were carried out. This does not mean that these nuclear extracts will necessarily contain factors that interact with DNA sequences of the maize *zrp2* promoter. However, these results clearly demonstrate that root and leaf nuclear extracts prepared as described above, contain functional nuclear factors, fully capable of specifically interacting with cognate DNA sequences.
CHAPTER 3: ANALYSIS OF THE DNA-PROTEIN INTERACTIONS IN THE PROXIMAL REGION OF THE ZRP2 PROMOTER

Introduction

ZRP2 (Zea Root Preferential) mRNA in maize has been found to be preferentially accumulated in roots (and stems) compared to the leaves, ears, silks and tassels (John, 1991; Held, 1993). The zrp2 cDNA clone was isolated by differential screening of a cDNA library constructed from poly (A)+ RNA, isolated from 9-d-old maize roots (John et al., 1992). The isolation and characterization of the near-full length cDNA clone and the genomic clone has been described in John (1991) and Held (1993) respectively. Analysis of mRNA distribution along the root has shown that it is undetectable in the first centimeter of the root containing the root tip, and is uniformly distributed at high levels along the rest of the root (Held, 1993). Cellular localization by in-situ hybridization studies has revealed that the ZRP2 mRNA is localized predominantly in the cortex of the root (Held et al., 1993). Comparison of the zrp2 nucleotide sequence with nucleotide sequences present in GenBank has shown no overall similarity. However, a stretch of 157 nucleotides between -209 and -366 in the promoter show 73% identity with the nucleotides of an intron sequence of a root-preferentially expressed α-tubulin from maize (Held, 1993).

Organ-preferential or specific accumulation of mRNA can be due to differential transcriptional and/or post-transcriptional (RNA processing, transport, degradation) regulation. Nuclear run-on experiments using nuclear extracts of different organs can be used to determine whether differential accumulation of transcripts is due to regulation at transcription. Nuclear run-on assays have shown dramatically higher levels of ZRP2
transcription in the nuclei isolated from young roots, compared to young leaves (Held, 1993). This result indicates that the ZRP2 mRNA levels are regulated primarily at the transcriptional level.

Different approaches have been used for analysis of the promoter region for cis-elements associated with organ-preferential/specific transcription. Promoter deletion/reporter-gene fusion constructs in transient or in stable-transformant expression studies can indicate the region of the promoter containing cis-elements important in modulating organ-preferential transcription. Analysis of promoter regions for DNA-protein interactions, using nuclear extracts of different organs can reveal location of cis-elements on DNA, and the pattern of the interaction. A correlation between the presence of a DNA-protein complex or the abundance of a binding protein, with the transcriptional activity of its target gene may provide circumstantial evidence for its involvement in the regulation of transcription (Forde, 1994; Cushman and Bohnert, 1992). A different pattern of protein-DNA interactions in different organs where transcription is differentially regulated may indicate specific cis-elements involved in organ-preferential/specific transcription (Manzara et al., 1991). Several cis-elements responsible for organ-specific/preferential transcription have been identified by the presence of such a correlation between the pattern of DNA-protein interactions and the differential transcription of the mRNA between two organs (Manzara et al., 1991; de Pater et al., 1993; Meller et al., 1993).

In the present study, 1 kbp of the promoter region, upstream of the transcription start site was analyzed for DNA-protein interactions by gel-shift assays, to determine whether the differential transcriptional activity in roots and leaves correlates with organ-specific differences in DNA-protein interactions. The first kbp was chosen because most of the
reported cis-elements that are responsible for organ or root-specific transcription are located within this region. In the case of the tobacco root-specific gene *TobRB7*, sequences up to 636 base pairs upstream from the transcription initiation site were sufficient to direct root-specific expression of the reporter gene, β-Glucuronidase (GUS), in transgenic tobacco. A negative regulatory element was also located between 813 and 636 bp upstream of the transcription start site (Yamamoto et al., 1991). In maize *Adh2* gene, an element located at 160 base pairs 5' to the transcription initiation site acts as an activator in the meristem and vascular tissue of roots and in the vascular tissue of stems and leaves (Paul and Ferl, 1994). In the tomato anionic-peroxidase (*tapl*) promoter, the region up to -202 bp from the transcription start site contains cis-elements necessary for developmentally-regulated tissue-specific expression of GUS reporter-gene (Mohan et al., 1993). In general, DNA sequences 5' proximal to the transcription start site of different *rbcS* genes are sufficient to direct organ-specific expression of reporter genes in transgenic plants (for review, see Dean et al., 1989). In addition, the 1 kbp *zrp2* promoter has been shown to be capable of directing transcription in maize roots (Held, 1993).

A *zrp2* promoter fragment containing DNA sequences from -472 to -180 from the transcription start site interacted differentially with nuclear protein extracts from roots and leaves. Further analyses were carried out in an attempt to delimit and differentiate the binding sites corresponding to these DNA-protein complexes.
Materials and Methods

Growth and harvest of seedlings

Growth of seedlings and harvest of root and leaf material is described in chapter 2. To prepare nuclear extracts of leaf base and leaf blade, leaves were excised approximately at one centimeter from the base, for separation into blade and base segments.

Isolation of nuclei and nuclear protein extract preparation

Isolation of nuclei, preparation of crude nuclear extracts and determination of the concentration of the protein is as described in chapter 2.

Preparation of the probe and the competitor DNA

Sub-cloning of the first kbp of the zrp2 promoter  Manipulation of DNA was done essentially as described by Sambrook et al (1989). The first kbp of the zrp2 promoter has been cloned and sequenced (Held, 1993). This DNA fragment was cloned into EcoRV and SphI sites of pGem 1.0 (modified pGem -3z) resulting in pZ2ES (Fig. 3.1). The plasmid pGem 1.0 was constructed by cloning a custom polylinker into BamHI and HindIII sites of the pGem -3z (Held, 1993).

The plasmid pZ2ES was digested with SalI to excise a DNA fragment of 796 bp at the extreme 5' end up to the SalI site situated at -180 bp. Religation of the resultant plasmid yielded pZ2(205)SS, containing the most 3' region of the one kbp promoter from -180/SalI to +25/SphI (Fig. 3.2A).

Digestion of pZ2ES with NsiI produced two DNA fragments of 270 bp and 240 bp. Religation of the resultant plasmid yielded pZ2(497)NS containing DNA sequences from
Figure 3.1 The plasmid pZ2ES in pGem 1.0. The plasmid pGem 1.0 was constructed by cloning a custom polylinker (Held, 1993) into the polylinker of pGem -3z. The TATA box, transcription start site (+1), and restriction sites used for dividing the 1 kbp zrp2 promoter are indicated.
-472 to +25 (Fig. 3.2B). The excised 270 bp DNA fragment contained sequences from -742 to -472, and the 240 bp DNA fragment contained sequences from the most 5' end of the 1 kbp promoter to -742. These DNA fragments were cloned into pGem 1.0, linearized with NsiI, to produce pZ2(270)NN and pZ2(240)NN respectively (Fig. 3.2C and 3.2D).

**DNA fragments spanning the junctions** The plasmid pZ2ES was digested with EcoRV and SphI to excise the entire 1 kbp promoter, which was gel-purified. Digestion of this DNA segment with Hinfl generated a 258 bp fragment (-938 to -680) spanning the -472/NsiI site, a 266 bp fragment (-645 to -379) spanning the -472/NsiI site and a 122 bp fragment (-257 to -135) spanning the -180/SalI site, among others. Digestion of pZ2.1EATG in pBluescript SK (Fig. 3.3), which contains the zrp2 sequence up to +50 (Held, 1993), with SalI and XhoI generated a 230 bp fragment spanning the +25/SphI site.

**Sub-division of the -472/NsiI to -180/SalI DNA fragment** The plasmid pZ2(497)NS was double-digested, with XbaI/NheI and XbaI/TaqI separately, to generate three smaller fragments spanning the entire length of the NsiI-SalI DNA fragment. Digestion with XbaI and NheI yielded a 97 bp XbaI to -379/NheI fragment. Digestion with NheI and TaqI yielded a 121 bp -379/NheI to -258/TaqI, and a 78 bp -258/TaqI to -180/TaqI fragment (Fig. 3.4). These fragments were separated by electrophoresis on 4% NuSieve-agarose gels and were purified by using Gene Clean(Gene Clean II- Bio101).

**Deletion of protein-binding region and cloning** The pZ2(497)NS was digested with NsiI and SalI to excise a 292 bp fragment from -472/NsiI to -180/SalI. Both the fragment and the vector containing -180/SalI to +25/SphI region were gel purified by Gene Clean.
Figure 3.2 The plasmids containing 200 - 300 bp DNA fragments spanning the entire 1 kbp zrp2 promoter. A. The plasmid pZ2(205)SS in pGem 1.0. The TATA box, transcription start site (+1), and restriction sites used for dividing the 1 kbp zrp2 promoter are indicated. B. The plasmid pZ2(497)NS in pGem 1.0. The TATA box, transcription start site (+1), and restriction sites used for dividing the 1 kbp zrp2 promoter are indicated. C. The plasmid pZ2(270)NN in pGem 1.0. D. The plasmid pZ2(240)NN in pGem 1.0
Figure 3.3 The plasmid pZ2.1EATG in pBluescript SK. The TATA box, transcription start site (+1), and restriction sites used for dividing the 1 kbp zrp2 promoter are indicated.
Figure 3.4 Division of the 1 kbp zrp2 promoter to delimit the region containing protein-binding sites.
The -472/NsiI to -180/SalI fragment was digested again with TaqI, which generated three DNA fragments of 82 bp -472/NsiI to -394/TaqI, of 136 bp -394/TaqI to -258/TaqI, and of 78 bp -258/TaqI to -180/SalI. The 82 bp fragment, the 78 bp fragment, and the vector containing -180/SalI to +25/SphI were religated to form pZ2(497-136)NS in pGem1.0.

Sub-division of the -379/Nhel to -258/TaqI DNA fragment The DNA fragment of -379 to -258 was further divided into two by restriction digestion of the plasmid pZ2ES with HindII and Nhel, which generated a 48 bp -379/Nhel to -330/HindII fragment and a 73 bp -330/HindII to -257/HindII fragment, among other fragments (Fig. 3.4). These were isolated following end-labeling and electrophoresis on 12% non-denaturing polyacrylamide gels.

Cloning of the oligonucleotides spanning the -330 to -257 region Two oligonucleotides of 48 bp and 46 bp, containing 38 bp and a 36 bp of promoter sequence spanning the entire -330 to -257 (73 bp) region of the zrp2 promoter, were synthesized at the ISU Nucleic Acids Facility. They were synthesized as single-stranded oligonucleotides. Ten to twenty five micrograms of each were annealed by heating to 90°C for 10 min followed by slow cooling to room temperature. Nhel sites were included in both oligonucleotides to use as a unique restriction site to screen for transformants. The oligonucleotide of 48 bp was synthesized with BamHI and SalI sites at the ends (Fig. 3.5) and the 46 bp oligonucleotide was synthesized with BamHI and XbaI sites at the ends (Fig. 3.5). These restriction sites were included in such a way that when the two strands are annealed, it would leave intact overhangs of restriction sites to facilitate direct cloning into
Figure 3.5 Sequence of the annealed, synthetic oligonucleotides, A 48 bp and B 46 bp oligonucleotides containing 38 bp and 36 bp of the zrp2 promoter sequence, spanning the region of -330 to -257. Restriction sites at the ends and the unique restriction site included in the oligomer are indicated. The promoter sequences are shown in bold.
pBluescript. Annealed oligonucleotides were cloned into pBluescript KS, double-digested with BamHI/SalI and BamHI/XbaI, to generate pZ2(48)BS and pZ2(46)BX respectively.

Oligonucleotides 40 bases in length, containing the sequences at -315 to -287 spanning the junction between the 46 bp and the 48 bp DNA fragments were synthesized at the ISU Nucleic Acids Facility. They were synthesized as two single-stranded oligonucleotides, containing the NheI site as a unique restriction site and restriction sites BamHI and SalI at the ends. Single-stranded oligonucleotides were annealed following the same method mentioned above (Fig. 3.6).

**Preparation of the probe and the specific competitor DNA** Plasmids containing the specific zrp2 promoter sequences were double-digested with restriction enzymes that leave 5' overhangs at the ends, to use both as probes and as specific competitor DNA. The plasmid pZ2(205)SS (Fig. 3.2A) was digested with BamHI and SalI to generate a 207 bp fragment containing sequences from -180 to +25. Plasmids pZ2(497)NS, pZ2(270)NN, and pZ2(240)NN were digested with XbaI and SalI to generate promoter fragments of 293 bp, 278 bp and 248 bp respectively.

The plasmid pZ2ES (Fig. 3.2B) was digested with EcoRV and SphI to excise the entire 1 kbp promoter, which was gel-purified. Digestion of this DNA segment with HindIII and NheI generated 258 bp, 266 bp, and 122 bp DNA fragments spanning the junctions of the first divisions of the zrp2 promoter. Digestion of pZ2.1EATG in pBluescript SK with SalI and XhoI generated a 230 bp fragment spanning the SphI (+25) site.

The plasmid pZ2(497)NS was digested with XbaI and NheI to generate a 97 bp DNA fragment, and with XbaI and TaqI to generate 121 bp and 78 bp DNA fragments to span the entire length from -472 to -180 of the zrp2 promoter.
Figure 3.6 Sequence of the annealed, 40 bp synthetic oligonucleotide containing -315 to -287 of the zrp2 promoter. Restriction sites at the ends and the unique restriction site included in the oligomer are indicated. The promoter sequences are shown in bold.

The plasmid pZ2(497-136)NS containing a mutated promoter fragment was digested with XbaI and SalI to generate a 160 bp DNA fragment with a deleted protein-binding region.

The plasmid pZ2ES was first digested with EcoRV and SphI to excise the entire 1 kbp promoter. This promoter fragment was digested again with HinfI and Nhel to generate 48 bp and 73 bp DNA fragments spanning the length of -379 to -258, among other fragments.

The plasmid pZ2(48)BS was digested with BamHI and SalI, and the plasmid pZ2(46)BX was digested with BamHI and XbaI to generate DNA fragments of 48 bp and 46 bp respectively.
To obtain labeled probes, entire digestion reactions were end-labeled with the Klenow fragment of the DNA polymerase I, as described (Ausubel et al., 1994), using $^{32}$P-labeled dCTP or dATP depending on the nucleotides in the 5' overhang. For a 100 µl restriction digest reaction containing 10 µg of DNA, 4 µl of 5 mM dNTP mix (without the nucleotide of the label), 1 µl of the Klenow fragment (5,000 units/ml), and 100 µCi of the labeled nucleotide were added. After incubation for 20 min at room temperature, 4 µl of non-radioactive 5 mM nucleotide was added and incubated for 5 min at room temperature. The labeled DNA was precipitated by adding a half-volume of 7.5 M ammonium hydroxide and two volumes of 100% ethanol, followed by incubation at -20°C overnight. After gel-purification, the total incorporation of radioactivity was determined by liquid scintillation spectrometry.

**Preparation of a non-specific competitor DNA**  A non-specific competitor DNA was prepared by digestion of the plasmid pBluescript KS with appropriate restriction enzymes to generate DNA fragments containing sequences of the multiple-cloning site, in approximately the same length as the labeled probe. A non-specific competitor DNA of 292 bp was obtained by the restriction digestion of pBluescript KS with PvuII and XbaI. This fragment included most of the pBluescript multiple-cloning site.

**Purification of DNA**

Probe, specific competitor DNA, and non-specific competitor DNA fragments were gel-purified by electrophoresis in 12% non-denaturing polyacrylamide (29:1 acrylamide to bis-acrylamide) gels, or on 4% NuSieve-agarose gels. Polyacrylamide gels were electrophoresed in TBE (Tris-borate-EDTA) buffer at 80 volts for 1-1½ hours. Gel pieces
containing DNA were excised using an autoradiogram of the gel, as the template. Polyacrylamide gel pieces containing the probe, specific competitor DNA or non-specific competitor DNA, were placed in 50-100 µl of nuclear binding buffer (see materials and methods, Chapter 2) overnight, at 37°C to facilitate diffusion of the DNA into the buffer.

NuSieve-agarose gels were electrophoresed at 80 volts in TAE (Tris-Acetate-EDTA) buffer for an appropriate duration depending on the size of the DNA fragment. Gels were stained in ethidium bromide, and pieces of the gel containing DNA fragments were excised using a UV light-emitting transilluminator. DNA was extracted from the pieces of gel by using Gene Clean, and finally dissolved in nuclear binding buffer. Concentration of DNA was quantified by comparison to standards on ethidium bromide-stained agarose gels.

**Gel mobility-shift assay**

Gel mobility-shift assays were performed as in Singh et al. (1989), according to Garner & Revzin, (1981) and Fried & Crothers (1981). Binding reactions contained 10⁴ cpm of end-labeled probe, 1.0 µg of poly(dI-dC).poly(dI-dC), 2 µg of crude nuclear protein extract in nuclear binding buffer. Total volume of a reaction mixture was 20 µl. Reactions were begun by the addition of labeled probe to the reaction mixture containing poly(dI-dC).poly(dI-dC) and nuclear protein extract. After mixing, reactions were incubated for 30 min at 30°C.

In competition assays, competitor DNA was added to the reaction mixture and incubated at 30°C for 20 min, prior to the addition of labeled probe. Molar ratios from 1:1 to 1:200 of the probe to specific competitor (non-labeled probe) or non-specific competitor DNA were added in competition assays.
After incubation, reaction mixtures were electrophoresed in 4-7% non-denaturing polyacrylamide gels (80:1 acrylamide to bis-acrylamide) according to the size of the probe. Electrophoresis was done at 70 volts for 1-1½ hours (10V/cm) at room temperature, using 200 mM Tris, 220 mM Borate and 5 mM Na₂EDTA (pH 8.3) as the electrophoresis buffer. Following electrophoresis, gels were transferred onto 3M paper and dried at 80°C for 1 hr. Autoradiography was performed at -80°C with a single intensifying screen, or at room temperature overnight, without an intensifying screen.

Results

The nucleotide sequence of approximately one kbp of the putative promoter region, upstream of the transcription start site of the \textit{zrp2} gene, has been determined (Held, 1993). In order to investigate the interaction of nuclear proteins with regulatory elements that may be responsible for the root-preferential transcription, the \textit{zrp2} promoter was subjected to electrophoretic mobility-shift assays. The first kbp of the promoter from the transcription start site was subcloned into four fragments of 200 - 300 bp in length to use as probes in gel-shift assays (Fig. 3.7A). Crude nuclear protein extracts of 5-d-old leaves and 3-d-old roots were used for interaction with possible binding factors, as root-preferential transcription previously had been observed in the seedling of the same developmental stages (Held, 1993). All binding reactions contained duplex poly(dI-dC).poly(dI-dC) as a competitor for non-specific DNA-protein interactions.
DNA-protein interactions of the DNA fragments spanning the 1 kbp promoter

The promoter region most proximal to the coding region, (-180/Sall to +25/SphI), containing the TATA box and the transcription start site (-180/Sall to +25/SphI) showed one DNA-protein complex with both root and leaf nuclear protein extract, that was barely detectable (data not shown). Repetition of the experiment with varying experimental conditions using several concentrations of the probe, protein extract, poly(dI-dC),poly(dI-dC) and polyacrylamide did not improve the level of detection.

DNA-protein interactions were clearly observed with the promoter fragment from -472/NsiI to -180/Sall in gel-shift assays (Fig. 3.7B). A different binding pattern between leaf and root nuclear protein extracts was detected due to the differential mobility of DNA-protein complexes, as a result of the different molecular weights of the proteins involved in the complexes. The DNA-protein complexes were numbered according to the mobility of the complex. The Z2BP1 (zrp2-binding-protein 1) appears to be abundant with the nuclear protein extracts from both leaves and roots, while the Z2BP3 appears to be more abundant with the nuclear extract from leaves. The Z2BP2 complex is formed almost exclusively with the nuclear protein extracts from roots. The abundance of the complexes, proximity of the binding region to the transcription start site, and differential binding pattern with nuclear protein extracts from roots and leaves suggested that this region may play a role in controlling transcription in a root-preferential manner. Gel-shift assays performed with the promoter regions -742/NsiI to -472/NsiI and the most 5' end of the promoter to -742/NsiI (Fig. 3.7A) showed no interactions with protein extracts under the same experimental conditions (data not shown).

To determine whether the factor/s involved in the formation of complexes consisted of proteins, gel-shift assays were performed with boiled nuclear protein extracts (data not
Figure 3.7 Gel-shift analysis of the zrp2 promoter. A. Restriction map of the 1 kbp zrp2 promoter. Only restriction sites that are used to divide the promoter into 200-300 bp fragments are shown. Position of the TATA box, transcription start site (+1), and sizes of the fragments used as probes in gel-shift assays are indicated. The presence or absence of a gel-shifted DNA-protein complexes is indicated as + or -. B. Gel-shift analysis of 293 bp (-472/NsiI to -180/SalI) zrp2 promoter fragment. End-labeled 293 bp fragment (10,000 cpm) was incubated with (+) or without (-) 2 µg of nuclear protein extracts from leaves (+L) or roots (+R), in the presence of double stranded 1 µg of poly(dI-dC).poly(dI-dC). Z2BP1, Z2BP2 and Z2BP3 are DNA-protein complexes formed with factors in the nuclear protein extracts, according to their mobility. After electrophoresis, the dried gel was subjected to autoradiography for 45 min at -80°C.
shown). No complexes were observed indicating that the complexes were formed with protein factors. The experimental conditions were optimized for the best visualization of DNA-protein complexes. Figure 3.8 shows gel-shift assays with varying amounts of leaf nuclear protein extract and of DNA, as the probe. Two micrograms of leaf nuclear protein extract and 5 ng of DNA with 1 µg of poly(dI-dC).poly(dI-dC) in 20 µl reaction mixture were selected as optimum conditions (Fig. 3.8, lane 6). With the increase of nuclear protein, the amount of DNA-protein complexes increased to a point where the free probe was limiting (lanes 2-4, 5-7, and 8-10).

There could be additional protein-binding sites on the promoter, where the junctions of the above fragments are located. These sites could have been divided due to the above fragmentation of the promoter. To investigate this, DNA fragments spanning these junctions were obtained by restriction digestion of the 1 kbp promoter with Hinfl and Nhel (Fig. 3.9). A 230 bp fragment spanning the +25/SphI site was obtained by the digestion of pZ2.1EATG with Sall and XhoI. A 258 bp fragment spanning the -742/NsiI site, a 266 bp fragment spanning the -472/NsiI site, a 122 bp fragment spanning the -180/Sall site and the 230 bp fragment spanning the +25/SphI site were subjected to gel-shift assays. None of these fragments exhibited complexes with nuclear proteins (data not shown).

**Specificity of the DNA-protein interactions**

To test whether these DNA-protein interactions were specific, gel-shift assays were performed with the labeled -472/NsiI to -180/Sall DNA fragment in the presence of specific and non-specific competitor DNA. Figure 3.10 and 3.11 show competition assays with leaf and root nuclear protein extracts respectively. The non-labeled -472/NsiI to -180/Sall DNA fragment was used as the specific competitor DNA. As non-specific DNA,
protein (µg)   -     0.5    1.0    2.0    1.0    2.0    4.0    2.0    5.0    8.0
probe (ng)    2.0    2.0    2.0    2.0    5.0    5.0    5.0    10.0   10.0   10.0

1  2  3  4  5  6  7  8  9  10

**Figure 3.8** Gel-shift analysis with the 293 bp (-472/NsiI to -180/Sall) zrp2 promoter fragment with varying amounts of probe and leaf nuclear protein extract in the presence of 1 µg of double-stranded poly(dI-dC).poly(dI-dC). Two (lanes 1-4), five (lanes 5-7) and ten (lanes 8-10) nanograms of probe was incubated with 0.5 (lane 2), 1.0 (lane 3 & 5), 2.0 (lane 4, 6 & 8), 4.0 (lane 8), 5.0 (lane 9) and 8.0 (lane 10) µg of nuclear protein extracts from leaves. Five nanograms of probe with 2 µg of leaf nuclear protein extract was chosen as optimal conditions. After electrophoresis, the dried gel was subjected to autoradiography for 3½ h at -80°C.
Figure 3.9 Restriction enzyme Hinfl-generated DNA fragments to span the junctions of the promoter fragments used for gel-shift assays. Fragments spanning the junctions are indicated. The 230 bp DNA fragment which span the +25/SphI site is also indicated. The presence or absence of a gel-shifted DNA-protein complexes is indicated as + or -.
Figure 3.10  Competition assay using the 293 bp (-472/NsiI to -180/SalI) zrp2 promoter fragment and leaf nuclear protein extract. Gel-shift assays were performed with either free probe alone (lane 1), with no competitor (lane 2), with 1x, 2x, 5x, 10x, 20x, and 50x molar excess of cold specific competitor (lanes 3-8), and with 100x, 200x molar excess of non-specific competitor (lanes 9 & 10). After electrophoresis, the dried gel was subjected to autoradiography for 2 h at -80°C.
non-specific DNA - - - - - - - - 50x 100x
specific DNA - - 1x 2x 5x 10x 20x 50x - -
root protein - + + + + + + + +
labeled DNA + + + + + + + + + +
1 2 3 4 5 6 7 8 9 10

**Figure 3.11** Competition assay using 293 bp (-472/NsiI to -180/Sall) zrp2 promoter fragment and root nuclear protein extract. Gel-shift assays were performed with either free probe alone (lane 1), with no competitor (lane 2), with 1x, 2x, 5x, 10x, 20x, and 50x molar excess of cold specific competitor (lanes 3-8), and with 100x, 200x molar excess of non-specific competitor (lanes 9 & 10). After electrophoresis, the dried gel was subjected to autoradiography for 2 h at -80°C.
a non-labeled 292 bp DNA fragment from the pBluescript multiple-cloning site (PvuII - XbaI) was used. Molar ratios of 1:1, 1:2, 1:5, 1:10, 1:20, and 1:50 probe to specific competitor DNA (lanes 3-8), and molar ratios of 1:50 and 1:100 probe to non-specific competitor DNA (lanes 9 and 10), were used to compete with the probe in competition assays. Lanes 2 through 8 in figure 3.10 show a clear decrease in the formation of complexes Z2BP3 and Z2BP1 in the presence of increasing specific competitor in the reaction mixture. The Z2BP3 complex was fully competed away by 1:10 molar ratio of the specific competitor DNA (lane 6). The Z2BP1 complex was also competed away by 1:50 molar ratio of the specific competitor DNA (lane 8). Conversely, even a 1:100 molar ratio of non-specific competitor DNA was not able to compete with the probe for binding with protein (lanes 9 and 10). Failure to abolish or alter binding activity in the presence of molar excess quantities of non-specific competitor DNA of similar size shows that the interactions between the DNA and protein factors are specific.

Lanes 2 through 8 in figure 3.11 show a decrease in the abundance of the complexes Z2BP2 and Z2BP1 as they get competed away by the presence of increasing amounts of specific competitor DNA. Both complexes were competed away by 1:50 molar ratio of the specific competitor DNA (lane 8). Conversely, as with leaf nuclear protein extract, even a 1:100 molar ratio of non-specific competitor DNA was not able to compete with the probe for interactions with the protein (lanes 9 and 10). This shows that the interactions between the DNA and protein factors are specific.
Sub-division of the -472/NsiI to -180/SalI DNA fragment to delimit the protein-binding region

To further delimit the binding region, the -472/NsiI to -180/SalI fragment was divided into three smaller fragments. Restriction digest of pZ2(497)NS in pGem1.0 (Fig. 3.2B), with XbaI/Nhel and Nhel/TaqI, generated a 97 bp DNA fragment containing -472/NsiI to -379/Nhel, a 121 bp DNA fragment containing -379/Nhel to -258/TaqI, and a 78 bp DNA fragment containing -258/TaqI to -180/SalI/TaqI (Fig. 3.12). These DNA fragments spanning the entire -472/NsiI to -180/SalI region were subjected to gel-shift assays. The 97 bp and the 78 bp DNA fragments did not show any gel-shifted DNA-protein complexes (Fig. 3.13), whereas the 121 bp DNA fragment exhibited the formation of DNA-protein complexes with leaf and root nuclear protein extracts (lanes 5 and 6). This showed exactly the same pattern of binding as with the -472/NsiI to -180/SalI fragment, with Z2BP1 being common to both leaf and root nuclear protein extracts, Z2BP2 being detectable only with root nuclear extract and Z2BP3 being more abundant with the leaf nuclear extract. Gel-shift assay with -472/NsiI to -180/SalI region with deleted 121 bp protein-binding region exhibited no binding activity (Fig. 3.14). These observations provide evidence that the binding activity observed with the -472/NsiI to -180/SalI fragment is due to the interaction of nuclear proteins with the -379/Nhel to -258/TaqI region.

Analysis of the sequence of the 121 bp region revealed that this sector is AT-rich (79% A or T). AT-rich regions in promoters are known to interact with transcription factors in plants (Datta and Cashmore, 1989), in yeast (Struhl, 1985), and in mammals (Fashena et al., 1992). Moreover, this region contains a part of the sequence (extending from -366 to -209) that shows 72% nucleotide identity over a stretch of 159 bp of the first intron sequence of a root-preferentially expressed \( \alpha \)-tubulin gene from maize. Together these
Figure 3.12 Further separation of the binding region into smaller fragments. Restriction map of 293 bp -472/Nsil to -180/Sall zrp2 promoter region. Only restriction sites that are used to divide the region into a 97 bp, a 121 bp and a 78 bp fragments are shown. The presence or absence of a gel-shifted DNA-protein complexes is indicated as + or -. Region that show 72% nucleotide identity with the first intron sequence of the root-preferentially expressed α-tubulin gene, is shown as a hatched region.
Figure 3.13  Gel-shift analysis of the 97 bp (-472/NsiI to -379/NheI), 121 bp (-379/NheI to -258/TaqI) and 78 bp (-258/TaqI to -180/SalI), with nuclear protein extracts from leaves and roots. Lane 1 contain the probe only. Lanes 2 and 3 contain 10,000 cpm of probe, 2 µg of nuclear protein extracts from leaves or roots, and 1 µg of poly(dI-dC).poly(dI-dC). After electrophoresis, the dried gel was subjected to autoradiography at room temperature without an intensifying screen.
Figure 3.14 Gel-shift analysis of the mutated -472/NsiI to -180/Sall region (with deleted 121 bp protein-binding region). All lanes contain 10,000 cpm of probe, 2 µg of nuclear protein extracts from leaves or roots, and 1 µg of poly(dI-dC).poly(dI-dC). After electrophoresis, the dried gel was subjected to autoradiography for 3 d at -80°C.
observations suggest that this region may be important in root-preferential transcription. The 121 bp -379/NheI to -258/TaqI fragment was further divided into a 48 bp -379/NheI to -330/Hinfl, and a 73 bp -330/Hinfl to -258/TaqI (Fig. 3.15). Restriction digest of the 1 kbp promoter with NheI and Hinfl generated these fragments among others. Gel-shift assays with the 48 bp fragment showed no binding activity (data not shown), whereas the 73 bp fragment showed interaction with nuclear protein that led to the formation of all three complexes (Fig. 3.16). As with 121 bp fragment, the 73 bp fragment also showed Z2BP1 with both leaf and root nuclear protein extracts, Z2BP2 predominantly with the root nuclear protein extract, and Z2BP3 predominantly with leaf nuclear protein extract.

In particle-bombardment-mediated transient-expression assays using the 1 kbp promoter, a significant amount of transcriptional activity was observed in the base of leaves and roots, and very low activity in leaf blades (described in Chapter 5). To investigate whether the DNA-binding activity is different between leaf blade and leaf base, and to see whether any binding activity is similar in leaf bases and roots, gel-shift assays were performed with nuclear protein extracts of leaf blades and leaf bases. Five-day-old leaves were excised approximately one centimeter from the base of the leaf, separating the organ into leaf base and blade areas for separate extraction of nuclear proteins. Gel-shift assays with the 73 bp DNA fragment from -330/Hinfl to -258/TaqI showed that the binding activity of leaf base and leaf blade is different (Fig. 3.17). The complex Z2BP1 is more abundant with the nuclear protein extract of leaf base than with leaf blade, and Z2BP3 is more abundant with nuclear protein extracts of leaf blade than with leaf base (compare lanes 4 and 5). Apparently, Z2BP1 with the whole leaf nuclear extract is contributed primarily by the leaf base, and Z2BP3 is mostly contributed by the leaf blade.
Figure 3.15 Further separation of the 121 bp binding region into a 48 bp and a 73 bp fragments. Only restriction sites that are used to divide the region into a 97 bp, a 121 bp and a 78 bp fragments are shown. The presence or absence of a gel-shifted DNA-protein complexes is indicated as + or -. Hatched region indicate the sequence that has 72% nucleotide identity with the first intron sequence of the root-preferentially expressed α-tubulin gene in maize.
Figure 3.16 Gel-shift analysis of the 73 bp (-330/Hinfl to -258/TaqI) region, with nuclear protein extracts from leaves and roots. Lane 1 contains the probe only. Lanes 2 and 3 contain 10,000 cpm of probe, 2 µg of nuclear protein extracts from leaves or roots, and 1 µg of poly(dI-dC),poly(dI-dC). After electrophoresis, the dried gel was subjected to autoradiography for 5 h at -80°C.
Figure 3.17 Gel-shift analysis of the 73 bp (-330/Hinfl to -258/TaqI) region, with nuclear protein extracts from whole leaves (+L), roots (+R), leaf bases (+L bs) and leaf blades (+L bld). Lane 1 contains the probe only. All other lanes contain 10,000 cpm of probe, 2 µg of nuclear protein extracts from leaves or roots, and 1 µg of poly(dI-dC).poly(dI-dC). After electrophoresis, the dried gel was subjected to autoradiography for 5 h at -80°C.
The higher molecular weight complex formed with the nuclear protein extract from the leaf base migrates slightly more slowly than Z2BP3 in leaf blade or whole leaf extracts (lanes 2, 4 and 5). The complex Z2BP1 with the nuclear protein extract from leaf base is as abundant as with nuclear protein extract from roots (lanes 3 and 4). Abundance of this complex correlates with the higher transcriptional activity seen in leaf bases and in roots (Chapter 5), suggesting that this complex may play a role in transcriptional regulation.

Analysis of the sequence of the 73 bp region revealed the presence of three 8 bp repeats of an AT-rich element, ATTTTTAT. This element is similar to the AT-1 element found in the promoters of several light-regulated photosynthetic genes including cab and rbcS of tomato, pea and tobacco (Datta and Cashmore, 1989). Single-stranded oligonucleotides were synthesized, annealed and cloned to generate DNA fragments 36 bp and 38 bp, spanning the 73 bp region (Fig. 3.18A). The oligonucleotides were made so that these three repeats are contained in one DNA fragment (in the 38 bp fragment). Gel-shift assays were performed to investigate the binding activity of the fragments. The 36 bp fragment exhibited no binding activity (data not shown), whereas the 38 bp fragment formed complexes with both root and leaf nuclear extracts (Fig. 3.18B). When compared with the 73 bp fragment, the intensities of the Z2BP3 formed is drastically decreased with the 38 bp fragment, indicating that the binding capacity is low. This may be due to loss of a part of the binding site or flanking regions of the binding site, the presence of which are important in forming an intact complex. With the 38 bp fragment, the Z2BP2 complex is undetectable with the root nuclear protein extract (lane 3). This could be due to the absence of the intact binding site on the 38 bp fragment. Moreover, absence of any complexes formed with the 36 bp fragment suggests that the binding site responsible for
Figure 3.18 Gel-shift analysis with the 36 bp and the 38 bp DNA fragments spanning the 73 bp (-330/Hinfl to -258/TaqI) region. A. Further separation of 73 bp region into two fragments of 36 bp and 38 bp. The 8 bp repeat sequences are indicated in arrows. B. Gel-shift assay with 38 bp (-298 to -258) region, with nuclear protein extracts from leaves and roots. Lane 1 contains the probe only. Lanes 2 and 3 contain 10,000 cpm of probe, 2 µg of nuclear protein extracts from leaves (+L), or roots (+R), and 1 µg of poly(dI-dC),poly(dI-dC). After electrophoresis, the dried gel was subjected to autoradiography for 7 h at -80°C.
the formation of Z2BP2 may be located in the junction of the two fragments. A careful search of the sequence of these fragments revealed the presence of another 9 bp AT-rich direct repeat located in the junction of the above two fragments and, therefore, had been divided in the center. To test whether the Z2BP2 binding site consisted of these two repeats, a 40 bp oligonucleotide spanning the junction of the above two fragments, containing the intact 9 bp direct repeat, was synthesized (Fig. 3.19). Gel-shift assays with the fragment failed to show any significant binding activity (data not shown). This result indicates that the intact binding site of the Z2BP2 complex does not consist only of the above two 9 bp repeats.

**Discussion**

Previous work from this laboratory has shown that ZRP2 mRNA is more abundant in roots than in leaves (John, 1991; Held, 1993). *In vitro* transcription run-on assays using nuclei isolated from 3-d-old roots and 5-d-old leaves has shown high levels of transcription in roots compared to undetectable levels in leaves (Held, 1993).

To define regions in the promoter, containing cis-elements responsible for root-preferential transcription, DNA-protein interactions of the first kbp of the zrp2 promoter were investigated. A 293 bp DNA fragment containing sequences from -472 to -180 showed distinct interactions with protein factors, while fragments containing sequences further upstream did not show any binding activity, in spite of the localized AT-rich (-843 to -752 and -738 to -659) and GC-rich (-546 to -525) regions in the promoter. The DNA fragment from -180 to +25, containing the TATA box, showed barely detectable binding activity with nuclear protein extracts of roots and leaves, in spite of obvious interactions
Figure 3.19 Location of AT-rich repeat sequences with regard to the synthesized oligonucleotides, which span the 73 bp promoter sequence (-330/Hinfl to -258/TaqI). The filled boxes indicate the 8 bp repeat sequences, TATTTTTA and the hatched boxes indicate the 9 bp repeat sequences, ATTTAAAA. The arrows indicate the orientation of the repeats.
with RNA polymerase and associated proteins. The experimental conditions used for isolating nuclear proteins and gel-shift assays may not be the optimal conditions for detection of these complexes.

Abundance of the complexes formed with -472 to -180 promoter fragment indicate the abundance of the interacting protein factors, in the nuclear protein extracts. The same DNA-protein interactions exhibited with this fragment were shown with the 121 bp (-379 to -258) fragment and the 73 bp (-330 to -258) fragment (Figures 3.13 and 3.16). This indicates that all the cis-elements interacting with nuclear factors are located in the 73 bp region.

The Z2BP1 complex is formed with both root and leaf nuclear protein extracts (Figures 3.7b, 3.15 and 3.16). Further analysis with nuclear protein extracts from leaf base, shows that the Z2BP1 with the whole-leaf nuclear extract is primarily contributed by the leaf base (Fig. 3.17). Mobility of DNA-protein complexes are effected mainly by the size of the bound protein (Ausubel et al., 1994). Complexes formed with protein factors of higher molecular weight are larger and migrate slower in the gel. When Z2BP2 and Z2BP3 are compared, the size of the protein(s) associated with the complex Z2BP3 is larger than that of Z2BP2. Moreover, Z2BP2 is formed almost exclusively with nuclear protein extracts of roots. Only a trace of Z2BP2 was able to be detected with nuclear protein extracts of leaf base (Fig. 3.17). This indicates that the interaction between promoter fragment and specific nuclear factor(s) that form Z2BP2 is more favored in roots. That could be due to the abundance of the active form of the specific factor in nuclear extracts of roots, compared to leaves. Conversely, formation of Z2BP3 is favored in leaves, probably due to the abundance of the active form of the specific nuclear factor/s in nuclear protein extracts of leaves. A correlation between the abundance of a binding
protein and the transcriptional activity of the gene may provide circumstantial evidence for its involvement in gene regulation (Forde, 1994). The above results clearly show that the interaction of nuclear factors with the same promoter fragment is different in roots and leaves, indicating existence of different mechanisms for transcriptional regulation of the \textit{zrp}2 gene in these organs.

The higher transcriptional activity of \textit{zrp}2 in roots correlates with the formation of Z2BP2, and the absence of Z2BP3, in roots. Similarly, low transcriptional activity in leaves correlates with the formation of Z2BP3, and the absence of Z2BP2, in leaves. Two models can be derived from these observations: 1) the interaction between the cis-element and the nuclear factor(s) that formed the Z2BP2 complex may activate transcription in roots; or 2) the interaction between the cis-element and the nuclear factor(s) that formed the Z2BP3 complex may repress transcription in leaves.

Gel-shift assays with 36 bp and 38 bp oligonucleotides spanning the 73 bp (-330 to -258) exhibited binding activity only with the 38 bp fragment, which consisted of the 3' region of the 73 bp fragment (Fig. 3.18A and B). Comparison of the intensities of the complexes show that the intensity of the Z2BP3 complex in the leaves has drastically decreased. This indicates loss of a sequence that facilitates binding, which may be situated in the periphery of the binding site. This may suggest that the binding site of the Z2BP3 may be located near the 5' end of the 38 bp fragment, or that a sequence located in the 36 bp fragment is associated with the binding of Z2BP3. It is possible for other sequences to be involved in binding cooperatively, thus facilitating the binding of a factor to a different cis-element.

Inability to detect Z2BP2 with either 38 bp or 36 bp fragments, lead to the hypothesis that the binding site may have been divided by the separation of 73 bp into two fragments.
Analysis of the sequence of the 73 bp fragment showed the presence of two 9 bp direct repeats 5' AATAATTTA 3', and three 8 bp repeats 5' TATTTTTTA 3' (two direct repeats and an inverted repeat, as in figure 3.19). The 3' end of the 9 bp repeat overlaps with the most 5' of the 8 bp repeat. Either of the 36 bp, or 38 bp fragment contained only one of the 9 bp direct repeats. However, gel-shift assays with a DNA fragment containing both repeats (40 bp) did not show any binding activity. This result suggests that the binding site of the Z2BP2 complex may be: 1) the 9 bp repeat, but flanking sequences are required for binding; or 2) the 8 bp repeat, but sequences in the 36 bp region are required for binding; or 3) sequences other than the repeats, but sequences of both 36 and 38 bp fragments are required for binding. Based on all the above observations, two possible models for the prominent interactions between the 73 bp protein binding region from the zrp2 promoter and protein factors from root nuclei, leaf base nuclei and leaf blade nuclei are illustrated in figure 3.20A and B.

The model 1 (Fig. 3.20A) is based on the possibility of the Z2BP2 binding site being the 9 bp repeats and the Z2BP3 binding site being located near the 5' end of the 38 bp DNA fragment. The model 2 (Fig. 3.20B) is based on the possibility of the Z2BP2 binding site being the three 8 bp repeats or a part of the three 8 bp repeats and the Z2BP3 binding site being located near the 5' end of the 38 bp DNA fragment. In roots both Z2BP1 and Z2BP2 complexes are abundant. In leaf blades Z2BP3 is abundant while Z2BP1 is very low. This trace amount of Z2BP1 in nuclear extracts from leaf blades may be due to contamination with leaf base region as a result of incomplete separation of leaf base and leaf blade. In leaf bases both Z2BP1 and Z2BP3 are abundant. The slightly higher molecular weight of the Z2BP3 (Z2BP3*) in leaf base nuclear extracts may be due to binding of both Z2BP3
Figure 3.20 Two possible models A and B, for the prominent interactions between the 73 bp fragment of the zrp2 promoter and protein factors from root, leaf base and leaf blade nuclear protein extracts. The DNA-protein complexes Z2BP1 and Z2BP2, were observed with nuclear proteins from roots, Z2BP1 and Z2BP3* complexes were observed with nuclear proteins from leaf bases and Z2BP3 was observed with nuclear proteins from leaf blades. The Z2BP1 was barely detectable with proteins from leaf blade nuclei.
B  Model 2

In roots:

\[ Z2BP2 \]

In leaf blades:

\[ Z2BP3 \]

In leaf bases:

\[ Z2BP3 \]

\[ Z2BP1 \]

Figure 3.20 continued.
and Z2BP1 (model 2), or due to a modification of Z2BP3, or due to association with another protein factor (model 1).

Based on the results of the transcription run-on assays (Held, 1993) the high transcription in roots correlates with the formation of Z2BP2 and the absence of Z2BP3, and undetectable levels of transcription in leaves correlate with the absence of Z2BP2 and the formation of Z2BP3. Based on these data it can be speculated that Z2BP2 may act as an enhancer and/or that Z2BP3 may act as a repressor.

Based on the results from the gel mobility-shift assays and transient-expression assays it can be proposed that the formation of the Z2BP1 complex may be responsible for the transcriptional activity in roots and leaf bases. The protein factor responsible for the formation of the Z2BP1 complex may be absent in leaf blades. Moreover, the protein factor(s) responsible for the formation of the Z2BP3 complex, which could be a repressor, is present in leaf blades. The very low transcriptional activity in leaf blades may be due to one or both of the above reasons. If the transcriptional activity is due to the formation of Z2BP1, it is possible that the transcriptional activity in roots and leaf bases are similar. The reason that the transcription run-on assays showed an undetectable levels of transcription (Held, 1993) could have been due to the dilution of ZRP2 transcripts by an excess of leaf blade nuclei compared to leaf base nuclei. However, the transient-expression assays of the promoter with protein-binding region deleted, exhibit a five-fold decrease in transcriptional activity in roots and an insignificant amount of decrease in transcriptional activity in leaf bases. This suggests that Z2BP2 may be involved in enhancing transcription in roots.

However, it is important to bear in mind that the binding of nuclear factors to cis-elements in gel-shift assays differs from the in vivo condition, since in gel-shift assays the specific cis-element is in excess in contrast to the binding factors. In the in vivo situations,
there can be competition between factors binding to the limited number of cis-elements, and thus, the binding affinities of the interactions and relative abundances of the factors undoubtedly plays an important role. Moreover, the chromatin structure in in vivo situation may influence the accessibility of the DNA to trans-acting factors, thus playing an important role in DNA-protein interactions.

The sequence of the first kbp of the zrp2 promoter is relatively AT-rich. Moreover, the sequence of the 73 bp, -330 to -258 region which contains all the above binding sites, is composed of repeat sequences which are predominantly consisted of A and T residues. The repeat sequences are composed entirely of A and T residues. In this respect, the 9 bp repeat sequences (putative Z2BP2 binding site) resemble the AT-rich sites located in the promoter regions of a number of plant genes, and these have been shown to bind nuclear factors. For example, conserved AT-rich sequences are present in the promoter region of the soybean lectin gene (Jofuku et al., 1987), of several nodule-related leghemoglobin genes (Jensen et al., 1988; Metz et al., 1988; Stougaard et al., 1987), of the nodule-enhanced glutamine synthatase gene from French bean (Forde et al., 1990) and of the soybean nodulin N23 gene (Jacobsen et al., 1990). Forde et al. (1990) have suggested that nuclear factors from a wide variety of plant organs that recognize cis-elements having the consensus sequence T/AAT$_{3-4}$A$_{1-2}$T may belong to an evolutionarily conserved and ubiquitous family. It is interesting to note that most of the above genes are expressed in roots.

Furthermore, conserved AT-rich sequence motifs present in the French bean phaseolin 5' flanking region are similar to AT-rich sequences present in 5' upstream region of several sunflower helianthinin genes (Bustos et al., 1989; Riggs et al., 1989). In the soybean glycinin A$_2$B$_{1a}$ gene promoter, an AT-rich sequence TAATAATTT (termed the glycinin
box) interacts with a nuclear factor from the embryo, which in turn may activate the seed-specific expression (Itoh et al., 1994) of the gene.

The 8 bp repeat TATTTTAT of zrp2 closely resembles the conserved AT-rich sequence having the consensus AATATTTTTATT (AT-1 box). Promoters of several light-regulated photosynthetic genes (rbcS and cab) from pea, tomato and tobacco also have AT-1 box sequences (Datta and Cashmore, 1989). Furthermore, from nuclear extracts of pea, Datta and Cashmore (1989) have demonstrated that the DNA-binding ability of AT-1 can be reversibly modulated by phosphorylation. AT-1 is active in the non-phosphorylated form and loses all binding ability as a result of phosphorylation. The effect of phosphorylation on binding abilities of nuclear factors that form complexes with zrp2 promoter has yet to be determined.

According to Forde (1994), there are two distinct types of nuclear proteins that interact with AT-rich upstream elements: the HMG-I-like factors (High Mobility Group I) and the apparently novel AT-rich Binding Factors (ATBF). The HMG proteins are a small (<30 kDa) heterogeneous group of eukaryotic chromosomal proteins, which are rich in acidic and basic amino acids and appear to be preferentially associated with transcriptionally active chromatin (Johns, 1982). Properties of HMG-like factors are fairly well defined based on studies in mammalian systems (Forde, 1994). Binding of these factors are not generally sequence-dependent and could tolerate considerable sequence-variety as long as a high AT content is maintained (Bustos et al., 1989). Conversely, ATBFs bind to their specific sequence. Most of the examples of AT-rich element binding factors mentioned above fall into this category.

Based on the criteria of Forde (1994), nuclear factors bound in Z2BP2 and Z2BP3 can clearly be categorized as ATBFs. Characteristics of ATBF-type nuclear factors include:
1) ability for binding in vitro to AT-rich sequences upstream of the TATA box (typically multiple sites 150 bp or more from the cap site); 2) high abundance (readily detectable in gel-shift assays); 3) heat sensitivity (unlike HMGIs); 4) multiple banding or broad banding in gel-shift assays (unlike HMGIs); 5) sequence specificity in binding, as demonstrated by competition experiments (unlike HMGIs); 6) target-sequence conforming to the motif A/TAT₃₋₅A₁₋₂T; 7) slow moving (unlike HMGIs).

Conversely, based on the observed characteristics from the gel-shift assays it is not possible to categorize the Z2BP1 complex definitively. The characteristics such as high mobility of the complex and inability to compete-out readily with specific DNA are shared with HMGI-like proteins. To be certain, the molecular weight of the protein factor has to be determined. Although the Z2BP1 complex cannot be fully competed away by the addition of 1:10 molar ratio of probe to specific competitor DNA (as with Z2BP2 and Z2BP3), it is competed away by 1:50 molar ratio of specific competitor DNA (Fig. 3.10). Furthermore, if Z2BP1 is formed due to interaction with HMGI-like protein, we should be able to observe it with any other AT-rich DNA fragment. But we were not able to observe it with AT-rich 9 bp repeat containing 36 bp fragment or 40 bp fragment. Based on these observations it is more likely that the factor responsible for Z2BP1 is also a ATBF-type protein.

A definitive role for AT-rich sequence elements in transcription of plant genes remains unclear. AT-rich cis-elements in light-regulated photosynthetic genes have been reported to act either as positive elements (Lam et al., 1990; Timko et al., 1985) or as negative elements (Castresana et al., 1988; Datta and Cashmore, 1989). Upstream AT-rich sequences that bind nuclear proteins have been shown to activate transcription in vivo (Bustos et al., 1989). The ATBFs associated with the diverse set of nodulin genes have
been implicated in enhancing transcription in nodules (Forde et al., 1990). However, most of the AT-rich binding sites characterized are frequently associated with positive regulatory elements (Forde, 1994).

According to Forde (1994), binding factors of ATBF-type are widespread among dicots, and evidence for their occurrence in monocots is limited up to now. Based on this, nuclear factors responsible for the complexes Z2BP3 and Z2BP2, and probably Z2BP1, are among the first few ATBF-type binding factors in monocots.

In conclusion, we have investigated the DNA-protein interactions of the first kbp of the zrp2 promoter proximal to the transcription start site. A differential binding pattern was observed with crude nuclear protein extracts from leaves and roots. These organ-specific differences in DNA-protein interaction correlate with the observed root-preferential transcriptional activity of the promoter, thus providing circumstantial evidence for its involvement in the regulation of transcription. The promoter fragment containing the binding sites was delimited to a 73 bp region, which contains two repeats of a 9 bp sequence and three repeats of a 8 bp sequence, all consisting only of A and T residues. Complexes Z2BP2 and Z2BP3 show similar binding properties to ATBF-type protein-DNA interactions. These could be among the first ATBF-type binding factors in monocots.
CHAPTER 4: DETERMINATION OF THE NUCLEOTIDE SEQUENCE OF
THE UPSTREAM REGION OF THE ZRP2 PROMOTER

Introduction

Transcription is regulated by interaction of multiple DNA-binding proteins with an array of cis-acting elements of the gene. These transcription regulatory cis-elements are primarily located in the upstream region of the gene. However, it is not impossible for cis-elements to be located in other regions, such as the 5' untranslated region, the coding region (Dietrich et al., 1992), the intron sequences, or in the 3' untranslated region of a gene. Cis-elements can be dispersed even as far as several kbp upstream from the transcription start site. These long-range cis-elements can affect transcription by tracking (Wang and Giaever, 1988), looping (Lobell and Schleif, 1990) or through protein bridges (Kim and Landy, 1992).

A genomic clone of the zrp2 gene has been isolated, and the transcription start site has been mapped. A total of 4.7 kbp of the upstream sequence has been cloned and the first kbp proximal to the transcriptional start site including the TATA box has previously been sequenced (Held, 1993).

Transcriptional analysis with the 1 kbp, 2 kbp and total 4.7 kbp (Fig. 4.1) promoter regions in the particle bombardment-mediated transient-expression assay has shown that the transcription activity increases with the length of the promoter in roots (Held, 1993). In leaves, the transcriptional activity was greatest with the 2 kbp promoter (Held, 1993). These data indicate the presence of other sequence elements located in the upstream region (upstream of 1 kbp) are involved in enhancing the transcriptional activity of the promoter.
Figure 4.1 The partial restriction map of the upstream region of the \textit{zrp2} gene up to the putative translation start site. The position of the TATA box, transcription start site, and 1 kbp, 2 kbp and the 4.7 kbp fragments of the promoter are indicated.
Determination of the sequence of the upstream region is necessary if we are to investigate DNA-protein interactions or to determine likely sequence elements (such as AT-rich regions, direct or inverted repeat sequences, etc.) that are important in DNA-protein interactions. The presence of the same sequence elements in the upstream region, that are involved in DNA-protein interactions of the 1 kbp promoter (Chapter 3), could reveal important information associated with the involvement of these cis-elements in regulation of transcription. Moreover, sequence comparison of the promoter may reveal similarities to other promoter sequences of genes in the GenBank, which could reveal common factors by which these genes are regulated.

The comparison of sequence of the 1 kbp region has shown 72% nucleotide identity with the first intron sequence of an α-tubulin gene in maize (Held, 1993). Transcripts of this α-tubulin gene preferentially accumulate in maize roots as do transcripts of zrp2. These observations indicate that this region may be important in root-preferential transcription. Investigation of DNA-protein interactions in the 1 kbp region showed differential pattern of the protein-binding in this region, with nuclear protein extracts of roots and of leaves (Chapter 3). In this study 1.6 kbp of additional sequence from the upstream region of the zrp2 promoter is analyzed.

Materials and Methods

Sub-cloning and sequencing of the -976/EcoRV to -1945/HindIII region

The upstream region from -976/EcoRV to -1945/Hind III of the zrp2 gene was sequenced at the ISU Nucleic Acids Facility according to the dideoxy-nucleotide chain termination method (Sanger et al., 1977), using double-stranded DNA templates (Chen and
Seeburg 1985). The subclone pZ2.1HE in pBluescript was used as the template for sequencing. The plasmid pZ2.1HE was constructed by digesting pZ2.45SS in pBluescript KS (Held, 1993) with EcoRV and HindIII, and cloning the resultant 1 kbp DNA fragment into pBluescript KS, digested with Smal and HindIII (Fig. 4.2).

The M13 forward and reverse primers were used initially to sequence from both ends of the insert. Synthesized 18-base oligonucleotides were used as primers as sequencing progressed. Primers used were, TT086 (5’ GACAAAGGGTCATCTCGC 3’), TT051 (5’ GATCTGGCTCGTGGCTAA 3’) and TT141 (5’ ATTTTGCTGATATGATC 3’) sequentially from the 5’ (HindIII) end of the promoter fragment and TT042 (5’ GAGGATGAATGAGGGTGT 3’), TT087 (5’ TTTGGAGGTCACGCTATC 3’) and TT050 (5’ GCGGGCGTCACGAC 3’) sequentially from the 3’ (EcoRV) end of the promoter fragment. Sequencing was done twice with each primer until the entire -976 to -1945 region was sequenced independently from both ends. The sequence was assembled using the UWGCG fragment assembly (version 7, April 1991) computer program from The Genetics Computer Group, Madison, Wisconsin.

Sub-cloning and sequencing of the -1945/HindIII to -2573/Apal region

The next 0.6 kbp (-1945 to -2573) of the zrp2 upstream region immediately 5’ to the above was sequenced using the same method at ISU Nucleic Acids Facility. The subclone pZ2.06AH in pBluescript KS was used as the template for sequencing. Initially the 2.6 kbp Apal to Apal fragment from pZ2.47SATG in pBluescript SK (Held, 1993) was cloned into pBluescript KS linearized with Apal, to form pZ2.26AA in pBluescript KS with the inserted promoter fragment in normal orientation. Restriction digestion of pZ2.26AA with
Figure 4.2 Construction of the plasmid pZ2.1HE in pBluescript KS. The plasmid pZ2.45SS in pBluescript KS (Held, 1993) was digested with EcoRV and HindIII, and the resultant 1 kbp DNA fragment was cloned into pBluescript KS, digested with SmaI and HindIII.
HindIII and religation resulted in generation of the plasmid pZ2.06AH in pBluescript KS (Fig. 4.3).

Initially, the M13 forward and reverse primers were used to sequence from both ends of the insert. Synthesized 18-base oligonucleotides were used as primers as sequencing progressed. Primers used were TT840 (5' TCTACTCTGTGAGTCTTA 3') and TT880 (5' CTTATGTCTGTGATGAT 3') sequentially from the 5' (ApaI) end of the promoter fragment and TT839 (5' ATCATCACAGACATAAGT 3') and TT881 (5' CTGAAGTCTCAACACTAA 3') sequentially from the 3' (HindIII) end of the promoter fragment. Sequencing was done twice with each primer until the entire -1945 to -2573 region was sequenced independently from both ends. The sequence was assembled using the UWGCG (version 7, April 1991) computer program for fragment assembly from The Genetics Computer Group, Madison, Wisconsin.

**Results**

The sequence of the upstream region of the zrp2 gene from -2573/ApaI to +25/SphI relative to the transcription start site, is shown in Figure 4.4. The sequence of the -976/EcoRV to +25/SphI region has been previously determined (Held, 1993). Analysis of the sequence showed localized AT- and GC-rich regions. Such AT-rich regions were found between -80 and -110, -215 and -328, -659 and -738, -752 and -843, -983 and -1049, -1173 and -1184, and -1313 and -1323. A GC-rich region is found between -546 and -525. The AT-rich stretch of nucleotides between -215 and -328 contained three 8 bp repeats (Fig. 4.4 bold and underlined) and two 9 bp repeats (Fig. 4.4 double underlined) consisting
Figure 4.3 Construction of the plasmid pZ2.06AH in pBluescript KS. Initially the 2.6 kbp ApaI-ApaI fragment from pZ2.47SATG in pBluescript SK (Held, 1993) was cloned into pBluescript KS linearized with ApaI, to form pZ2.26AA in pBluescript KS. Restriction digestion of pZ2.26AA with HindIII and religation, resulted in generation of the plasmid pZ2.06AH in pBluescript KS.
**Figure 4.4** The sequence of the entire 2.6 kbp *zrp2* promoter. The region from -976/EcoRV to +25/SphiI was sequenced by Held (1993). The region that shows 72% nucleotide identity with the first intron sequence of the root-preferentially transcribed α-tubulin gene from maize, (Montoliu et al., 1989) is underlined. The three 8 bp repeats (in bold) and the two 9 bp repeats (double underlined) in this region are as indicated. The other 8 bp repeat located at -1041 and -2053 is indicated in bold and is underlined. The TATA box (bold and underlined), the transcription start site (+1), and the putative translation start site (in bold) are also indicated.
GGGCCCCAAAG GAAGGAAGTA CTTTCAGTGG A TCAAGATGG TATGTTCCCT
-2523 GATGGGTATG CAGCTAACCT GAGTAGTTGG GGTGAACTTA TCTACTCTGT
-2473 GAGTCTTAGG GATGAAGAGT CATGACTTCC ACATATGGAT TGAACAGATT
-2423 CTTTCCTGTG CATGGACAAT CTGGGGCGGC ATCCAACAAC CCTCATGGAT
-2373 CGCCCGGCCA ATCGCCGCAC CAGTCCATCC GCCCCACTCG ATGAGACTTA
-2323 TGGTCTTAGT TGGAGACTT CAGAACTTTT TGATAATGCT GTATTGGATA
-2273 CTTATGTGTT TGTTGAGATC TTATGTGAGA ATTTTAGATAT TTATGTTTGT
-2223 ATGTCTTGTG TACTATGTGTT GTGTGGAGAT CATTTGGAAT TTATGTTTGT
-2173 GTTGAGATAT CAGGAATATC TATGTGGTAT CTGTTTCTTG TTTGGAGAAA
-2123 TATGTGACAT ATGTGATGTA TATGTGGGTAT CTTTTTGTTG TTTGGAGAAA
-2073 ATAGAGAAAG CAAAATAAAAA TGTGATATTT GGTCAGTTTGT GCGATGTTAA
-2023 CACTCGGCAA AAAGGTGCTT TGCCGAGTGT TAGGGCCATA GCACCTCGGTA
-1973 GAGAACAAT ACTTACCGCAC CGGTAAAGCT TTTTTGGCGA GTTTGTGGGC
-1923 CCTGGCCACTC AGCTTTGCGG AGTGCTCCAC ACAGGACTCG ACAAGAAACC
-1873 TGACAAATGG ACCCGCTGTT AAATCCCTTTA CCGAGTCGAG GTGATAGAC
-1823 ACTCGGCAA A GAATCTTCTGT TTCCGAGTGT CCGCTTAGAA CATTTGACAA
-1773 AGGGTCTATCT CCCCTACCGG GTGCTGTGAC GGGCCGCTTTT CTGAGGATGT
-1723 TGTAGGGCTGC TTTTCTGAGT ATTACACTTG AAAAAAAATG CTCGTGGCTA AGATCTCTCT TCTTCCCTTT
-1673 CTGACGTCTTC TTGTCAGGATG ATTACATCTC GCACGACTCG ATGAGGGAGGC ACGAGCCTGC
-1623 TTGAGTAGCT TTTCGAGTGT A TTTTACGCCC TCAGCAGGTTG TCGAGTTGAA
-1573 GGTAGTGACG GGGCTTTGCG AATAGTAAAA TCCAGCCCCT CCGCTGTGGG
-1523 AAAAACTGGA TGGATCTGGA CGTCGAGCTA AGATTCTCTT TCTTCCCTTT
-1473 GTAAAAAAAG AAAGGAAAAA AAAAAAGCT GCACGAGGGC CTGAGGCTGGT
-1423 GAGTAGGCTG ACTCGCAAA AAGGACCGCT GCATCGAGGATG ACGGAGGCTG
-1373 TATGATAGCTG ACCTCCAAAA TAAAGAAATC GGGCTGTGGC CCCGCGGTGC
-1323 AATATAATTTAA GGGCCACTGA TTAGGTGGGG TTAGGATAC ATCACTTCCAG
-1273 AACATCATCT GATCTGATCT GATTTGGTCT GATATGATCT GGGTAGTCAT
-1223 TTCTGCAATG AGCATCTATC AGGTGAACCA ATTAATATTG ATGACATTAT
-1173 GAGTTCCAAG ATATACTCTA AAGTGTTATC TAAATACAGA AGACATCTCGT
-1123 TCGTTCCTTG CCTATAAATC TAAAGGCTTT GTAAACCCT CATTCTACCCT
-1073 CTATATACGA AGACTCCTCT CTAATATAGTTAT CCGATTATGA TTTTTTTTAT
-1023 ATTTAGACAA TGGAAATAAAA TAGAACTAAA ATATATATAA GATGATATCT
- 973 GAGGACCCGA GATGGTAATG GGAGCTCGAT CCTGATTCT CCACGGAGAA
- 923 TTCCCTCTAGG ATATAGGTAA TTTGTCCTCA CGAGATTGGA AACCGGTAAT
- 873 TTTGGTCCCC ATGTGCCCCG CCCGCGAAGT TCTCTTGTAC TAAATTAGTCC
- 823 TATTTTCTATG TTTAAACTAT ACTAAAATT TAAATACAGA TCTATTAAAT
- 773 AATAGCAAAAC TAAATTCTAA AGTTGAATGC ACGGTCTGTT TTAATTCTGG
- 723 TTTGTCTCAAG TTATATTATG TTGATATAAT AAAAAATGAT TTGACCTTTTA
- 673 ATACGTATT TTTTCTCTAC GGGACGGAGT TCTCCAGGGA GATAAAACTCC
- 623 ATGATACAGA TGGGATGAAA GAAATAATCTC CCGTATGAAC TTTTCAGGA
- 573 ATGGGGATGG GCCAGAGAAA TTTTCTCTCT GGGGGACGGG GGGAGCCATA
- 523 TCTCTGGTGG AGAATTCCTA ATTATCATCC TTAATTTGGG TACATATATA
- 473 TGCTATATCT TTTTTTTTGT ACTGACATGT GGGAAAGTAT CCCATCTCAA
- 423 TAGTAGAAAA TCTGGGGAAC GGTAGGATCG AACACAAAGA TCAGCTAGCT
- 373 TGCTACCCAG GACGCATATA GCTAGAGGTT AATAATATCAGAT GAACTCAAAT
- 323 TTTTTTTTCT AAATATTAAA GGCTCTAAAAT TATTTTTTAT TAAAAAAAATA
- 273 ATAAAAATAT AGTTGACACC TTAATAATATA TATGTAATAC ATTTAAAAATGTC
- 223 TATTATTACC CCTACTTTATT GAGTTATGGT TCAGTTCTTG TCAGACGGAGA
- 173 GTAATGAGAT ATAGAATAAG GTACCCTATA GAATAAAGAA TCTTTCTCTG
- 123 AAAAGTCTGA CGTACGGTTAA TAAAGATACAA TAAAAAAAT ACAAAGAGAA
- 73 GCGCTGGACT GGAGATGCTC CTATATGCGG CAATGCCTGT GCTTATAAAT
- 23 AGCCACCTCG GTCGGCAAGG ACGTGAAACG CGAGCGCAGT GTGCAGTC

Figure 4.4 continued.
only of A and T nucleotides. The same 8 bp motif was also found at -1041 and -2052 (Fig. 4.4 bold and underlined). In addition, there are several sequence repeats found in the region between -2100 and -2330. These include five 9 bp repeats (ACTTATGTGTT) and three 14 bp repeats (TTATGTTTGTGTGTTG), which overlap with each other, and two 18 bp repeats (ATATGTGATGTATATGTG) and five 7 bp repeats (TATATGT), which overlap with each other.

**Discussion**

The zrp2 upstream region contained several AT-rich regions over the entire 2.6 kbp region. AT-rich regions are known to interact with protein factors to cause transcriptional regulation in plants (Lam et al., 1990), in yeast (Struhl, 1985), and in mammals (Fashena et al., 1992). In plants, AT-rich binding sites are frequently associated with positive regulatory elements (Forde, 1994). Gel-shift assays with the 1 kbp promoter region showed DNA-protein interactions only with the AT-rich region between -330 and -258 (Chapter 3). This region is spanned by two 9 bp direct repeats 5' AATAATT TA 3' and three 8 bp repeats 5' TATTTTTTA 3' (two direct repeats and the other as an inverted repeat) as shown in Figure 3.19. Direct repeats or inverted repeats frequently act as binding sites for protein dimers. The presence of these repeated motifs and their protein-binding activity strongly suggest that this region may be important in regulation of transcription.

Presence of the same 8 bp motif in the region upstream of the 1 kbp promoter suggests the possible involvement of these motifs as long-range effectors in transcriptional regulation. They could function as redundant elements or act synergistically in concert.
with the other three repeats located in the -330 to -258 region. If this motif acts as an enhancer, the latter may explain the increase of transcriptional activity with increasing length of promoter in the particle bombardment-mediated transient assays in roots.

The presence of several other repeated sequences in the ApaI/-2573 to HindIII/-1945 region, may also be important in regulation of transcription. The importance of these sequences could be determined by 'loss-of-function' and/or 'gain-of-function' studies.

Comparison of the total 2.6 kbp zrp2 upstream sequence with the sequences in the GenBank showed sequence similarities with upstream sequences of a few other genes. The region between -1915 and -1564 in reverse, showed 66.9% sequence identity with the upstream sequence of a zein gene (maize 19 kD protein; MZEZE19A- Accession # J01243) from maize. The region between -1720 and -1553 showed a 79.5% sequence-identity with the upstream sequence of a globulin gene (Maize globulin-1 gene, MZEGLOBO1 - Accession # L22344 or L22295) from maize. The zeins and globulins are storage proteins. The globulins are the major storage protein in maize embryos accounting for 10-20% of the embryo protein (Kriz, 1989). Sequence similarity for both of these storage protein genes lies between -1913 and -1553, with -1720 to -1564 portion being common to both. This may suggest the production of a storage protein as a possible function of the zrp2 gene.

Finally, the region between -2561 and -2514 (147 bp), which is very close to the most 5' end of the 2.6 kbp zrp2 upstream region, showed 60% nucleotide identity with the maize autonomous transposable element En-1, mosaic-protein gene (MZETNENSPM - Accession # M25427). The significance of this nucleotide sequence similarity is unclear.
CHAPTER 5: EFFECT OF THE PROTEIN-BINDING SEQUENCES ON TRANSCRIPTION OF THE ZRP2 PROMOTER

Introduction

The functional significance of promoter regions on the transcription of genes can be analyzed by two approaches. First, they can be analyzed for loss-of-function, where the activity of the promoter with or without the region of interest (5' end or internal promoter deletions) is compared. Second, they can be analyzed for gain-of-function, where the promoter region of interest is placed upstream of a heterologous minimal promoter and activity is compared. In both approaches a reporter gene is transcriptionally fused to the mutated or chimeric promoter, and the level of transcription is assayed.

In plants, *in vivo* assay systems such as transient-expression assay systems or stable transformation assay systems are widely used in analysis of promoters, as compared to *in vitro* transcription assay systems. Only a very few attempts have been made to investigate transcription of plant genes in *in vitro* transcription systems (Arias et al., 1993; Cooke and Penon, 1990), mainly due to the lack of a well-defined *in vitro* transcription system. Moreover, use of a homologous *in vivo* assay system have the advantage of the promoter being in its native context. Between transient and stable transformation systems, the main advantage of the transient system is the ability to obtain results in very few days compared to the time needed for transgenic plants to regenerate. Furthermore, the stable transformation of monocots has proven to be difficult (Gordon-Kamm et al., 1991), compared to dicots where a well-defined *Agrobacterium*-mediated DNA transfer system has been developed and can be applied.
Gene-transfer systems

There are direct and indirect systems for transfer of genes into plant cells. The bacterium *Agrobacterium tumefaciens* is the gene-transfer intermediate most frequently used in indirect gene transfer, for higher plants. The bacterium is naturally capable of transferring a piece of DNA into the genome of most dicotyledonous plants (Fraley et al., 1986; Bevan, 1984). The natural host-range of *Agrobacterium* is limited to dicots, and reports of *Agrobacterium*-mediated gene transfer into monocots are few (reviewed in Hooykaas, 1989).

Commonly used direct transfer methods include polyethylene glycol (PEG)-mediated or calcium-mediated uptake of DNA, or electroporation of naked DNA, into protoplasts or cell suspension cultures (Giovinazzo et al., 1992), micro-injection of DNA into embryos (Neuhaus et al., 1987) or protoplasts (Reich et al., 1987), and DNA-coated microparticle bombardment of intact cells (Klein et al., 1988a and b). Direct transfer methods can be used for stable or transient introduction of DNA, and are instrumental in stable introduction of foreign DNA into monocots (D'Halluin et al., 1992). Transgenic plants are generated at low frequencies following direct transfer of foreign DNA into intact cells by electroporation or microparticle bombardment (reviewed in Potrykus, 1990).

Transformation by microprojectile bombardment

Bombardment with microprojectiles has become one of the major techniques for the transformation of plant cells. This technique does not suffer from many of the restrictions characteristic of other techniques. Notably, there is no host range limitation, and bombardment is not restricted to certain specific tissues or cell types. Apart from that, this technique is much easier and less time-consuming to perform than the tedious task of
microinjecting embryos. Particle bombardment is the only available technique for direct transfer of genes where the cell wall need not be considered as an obstacle (Hamilton et al., 1992). For these reasons this is the most convenient technique applicable to species such as maize (Gordon-Kamm et al., 1990), sugarcane (Bower and Birch, 1992), rice (Christou et al., 1991) and wheat (Vasil et al., 1992; Weeks et al., 1993).

However, this method has a number of limiting factors as well. Apart from the technical considerations including the type of particle gun and experimental set-up, some tissues may be resistant to particle penetration due to a strong cuticle, lignified cell walls or a surface covered with trichomes. Moreover, as only a small fraction of the cell population is penetrated by a particle, efficiency of transformation is low.

**Parameters/function of the PDS-1000/He particle gun**

In this study the helium-driven PDS-1000/He particle gun was used to bombard DNA-coated tungsten particles into seedling roots and leaves of maize. A diagram of the set-up is shown in Figure 5.1. The efficiency of bombardment is determined by the nature and size of the microparticles, and by several other instrument parameters. The choice of the rupture disk which is designed to rupture at a particular pressure, determines the pressure of the helium that enters the bombardment chamber, which in turn determines the velocity to which the microparticles are accelerated. When the rupture disk bursts, the helium shock wave drives the macrocarrier, carrying DNA-coated microcarriers on its front surface, towards the stopping screen. The stopping screen retains the macrocarrier while allowing the coated particles to pass through to penetrate the target cells at the bottom of the bombardment chamber.
Figure 5.1 Schematic representation of the PDS-1000/He system (not to scale). X, Y and Z are the adjustable distances that influence the velocity of the microcarriers that hit the target cells. X. Rupture disk-macrocarrier gap distance. Y. Macrocarrier travel distance. Z. Target distance.
Analysis of the zrp2 promoter by particle bombardment-mediated transient assay

Analysis of the first kbp of the zrp2 promoter for DNA-protein interactions that may be involved in root-preferential transcription revealed that the region -379 to -258, relative to the transcription start site, interacts differentially with nuclear protein extracts of roots and leaves (Chapter 3). Gel-shift analysis with this region showed a DNA-protein complex common with both root and leaf (base) nuclear protein extracts (Z2BP1), a complex predominantly formed with root nuclear protein extract (Z2BP2), and a complex predominantly formed with leaf nuclear extract (Z2BP3).

To investigate whether cis-acting elements contained in this region are responsible for conferring root-preferential transcription, an internal deletion of the promoter including this region was tested for transcriptional ability. Transcriptional activities of the 1 kbp and the total 4.7 kbp promoter, with or without the internal deletion, were compared in microprojectile bombardment-mediated transient expression assays.

Materials and Methods

Growth of seedlings

Maize (Zea mays L. cv NKH31) seeds were surface sterilized in a 50% commercial bleach solution for ten minutes, rinsed three times with sterile water and imbibed for one hour. Approximately 50 seeds were placed on sterilized germination paper, which was then rolled and placed vertically in a polyethylene container containing 5 to 8 inches of sterile water. Seedlings were grown at 30°C under a 12 h light/12 h dark cycle in a growth chamber (as in Held, 1993).
Harvest and preparation of plant material for bombardment

Roots from three-day-old seedlings and leaves from five-day-old seedlings were harvested for bombardment. Approximately 3 cm of the terminal part of roots were excised with a razor blade. Primary leaves were harvested by gradually pulling the leaf from the sheath (as in Held, 1993).

Approximately 10 to 15 roots or 5 to 10 leaves were placed on sterile 1% water agar plates 2 inches in diameter, lined with sterile Whatman filter paper moistened with sterile distilled water. They were aligned in parallel at approximately equal distances apart, alternating in opposite orientations (Fig. 5.2). Plant materials were prepared for bombardment just prior to precipitation of DNA on microparticles.

DNA constructs used for bombardment

Six different promoter/GUS constructs were used in the transient expression assays: 1) p1068, 2) p-35S1068, 3) pZ2.1EATG in 1068, 4) pZ2.1EATG(-136) in 1068, 5) pZ2.47SATG in 1068 and 6) pZ2.47SATG(-136) in 1068 (Fig. 5.3). The plasmid p1068, containing the 35S promoter, the intron 6 sequence of alcohol dehydrogenase (INV6), the β-glucuronidase (GUS) coding sequence and the Nopaline synthase terminator (NOS) sequence in transcription fusion, were obtained from Sandoz/Northup King Corporation. The promoterless construct (-35S) was constructed by restriction digestion of p1068 with SacI to excise the 35S promoter, followed by religation. The plasmid pZ2.1EATG in 1068 was constructed by digesting out the SacI to XhoI fragment from pZ2.1EATG in pBluescript SK, which contains the first kbp of the zrp2 promoter up to the putative translation start site (-976/EcoRV to +62/ATG), and by cloning the excised
Figure 5.2 Five-day-old leaves and three-day-old roots of maize, prepared for bombardment with DNA-coated tungsten particles.
Figure 5.3 Promoter/GUS fusion constructs used in the particle bombardment-mediated transient-expression assay. A 136 bp region containing the protein-binding sequences was removed from the 1 kbp (1 kbp - 136 bp) and the 4.7 kbp (4.7 kbp - 136 bp) zrp2 promoter.
fragment into the XhoI- and SacI-digested p1068 (Held, 1993). Similarly the plasmid pZ2.47SATG in 1068 was constructed by digesting out the SacI to XhoI fragment from pZ2.47SATG in pBluescript SK, which contains the total 4.7 kbp zrp2 promoter up to the putative translation start site (SalI to +62/ATG) and cloning the excised fragment into the XhoI- and SacI-digested p1068. As the result both pZ2.1EATG in p1068 and pZ2.47SATG in p1068 were generated, which contained 1 kbp and 4.7 kbp of the zrp2 promoter respectively (Held, 1993).

Plasmids pZ2.1(-136)EATG in p1068 and pZ2.47(-136)SATG in p1068, containing 1 kbp and 4.7 kbp zrp2 promoter sequences with an internal deletion removing the protein-binding (as observed in Chapter 3) region, were obtained through the following steps.

First, the plasmid pZ2(497-136) in pGem 1.0, containing -472/NsiI to +25/SphI region of the zrp2 promoter with protein-binding sequences deleted, was constructed. To delete the 136 bp -394/TaqI to -258/TaqI region containing all the observed protein-binding activity, the plasmid pZ2(497)NS (Fig. 3.2B) was digested with NsiI and SalI to remove a 292 bp fragment from -472/NsiI to -180/SalI. Both the fragment and the vector containing the -180/SalI to +25/SphI region were gel purified by Gene Clean. The -472/NsiI to -180/SalI fragment was digested again with TaqI, which generated three DNA fragments of 82 bp (-472/NsiI to -394/TaqI), 136 bp (-394/TaqI to -258/TaqI), and a 78 bp (-258/TaqI to -180/SalI). The 82 bp fragment, 78 bp fragment and the vector containing -180/SalI to +25/SphI were religated to form pZ2(497-136)NS in pGem 1.0 (Fig 5.4).

Secondly, the -472/NsiI to +25/SphI unmutated region was replaced with the mutated -472/NsiI to +25/SphI in the plasmid pZ2ES in pGem 1.0, generating the plasmid pZ2(-136)ES in pGem 1.0. To replace unmutated -448/AflIII to +25/SphI of pZ2ES in pGem 1.0 with the mutated AflIII to SphI region, the plasmid pZ2(497-136)NS in pGem 1.0 was
Figure 5.4 Construction of the plasmid pZ2(497-136)NS in pGem 1.0. This plasmid was generated by deleting the protein-binding region from the plasmid pZ2(497)NS in pGem 1.0. The flanking regions of the deleted sequence is indicated by hatched area.
first subjected to restriction digestion by AflIII, which cut the plasmid in two places. One of the products, containing the mutated AflIII to SphI and some vector sequences, was isolated and subjected to digestion by SphI, which generated an AflIII to SphI fragment containing the mutated sequence. Meanwhile, the plasmid pZ2ES in pGem 1.0 was partially digested with AflIII by modifying the digestion conditions to obtain predominantly AflIII-linearized plasmid. Isolated linearized plasmid was digested with SphI which yielded a 1336 bp, a 2290 bp, a 3286 bp and a 472 bp fragment. The 3286 bp fragment containing the sequence of the promoter from the most 5' end to the -448/AflIII, was isolated and ligated with the above AflIII-SphI fragment with deleted protein-binding sequence, to form pZ2(-136)ES in pGem 1.0 (Fig. 5.5).

Thirdly, the -976/EcoRV to +25/SphI region in pZ2.1EATG in pBluescript SK, was replaced with the mutated -976/EcoRV to +25/SphI, to generate pZ2.1(-136)EATG in pBluescript SK. Similarly, the -448/AflIII to +25/SphI region in pZ2.47SATG in pBluescript SK, was replaced with the mutated -448/AflIII to +25/SphI to generate pZ2.47(-136)SATG in pBluescript SK. To replace the intact EcoRV to SphI region in pZ2.1EATG in pBluescript SK with the EcoRV to SphI fragment from pZ2(-136)ES containing the mutated sequence, both plasmids were digested with EcoRV and SphI individually. The resultant pBluescript SK vector sequence containing SphI to XhoI of the promoter was isolated and ligated with the mutated EcoRV to SphI fragment to form pZ2.1(-136)EATG in pBluescript SK (Fig. 5.6). To replace the intact -448/AflIII to +25/SphI of pZ2.47SATG in pBluescript SK by the mutated AflIII to SphI, the plasmid pZ2.47SATG in pBluescript SK was first subjected to partial restriction digestion by AflIII. To obtain predominantly AflIII-linearized plasmid, the digestion conditions were
Figure 5.5 Construction of the plasmid pZ2(-136)ES in pGem 1.0. This plasmid was generated by replacing the intact -448/AflIII to +25/SphI region of pZ2ES in pGem 1.0, by the mutated -448/AflIII to +25/SphI region from pZ2(497-136)NS in pGem 1.0.
Figure 5.6 Construction of the plasmid pZ2.1(-136)EATG in pBluescript SK. This plasmid was generated by replacing the intact -976/EcoRV to +25/SphI region of pZ2.1EATG in pBluescript SK, by the mutated -976/EcoRV to +25/SphI region from pZ2(-136)ES in pGem 1.0.
adjusted. Isolated 7.7 kbp linearized plasmid was digested with SphI which yielded a 7.4 kbp, a 5.1 kbp, a 2.6 kbp and a 472 bp fragment. The 7.4 kbp fragment containing the sequence of the most of the promoter, was isolated and ligated with the mutated AflIII-SphI fragment from pZ2(-136)E, to form pZ2.47(-136)SATG in pBluescript SK (Fig. 5.7).

Finally, the 35S promoter in p1068 was replaced with mutated 1 kbp and 4.7 kbp zrp2 promoter fragments to generate pZ2.1(-136)EATG in 1068 and pZ2.47(-136)SATG in 1068 respectively. To replace 35S promoter with mutated zrp2 promoters, both Z2.1(-136)EATG in pBluescript SK and pZ2.47(-136)SATG in pBluescript SK were digested with SacI and XhoI, and the excised mutated promoter fragments were cloned into SacI- and XhoI-digested p1068, to generate Z2.1(-136)EATG in p1068 and pZ2.47(-136)SATG in p1068 (Fig. 5.3).

**Preparation of DNA-coated microcarriers**

The M17 tungsten particles were used as microcarriers. Plasmids containing the 1 kbp and the 4.7 kbp zrp2 promoter fragments with or without the internal deletion, along with 35S and -35S promoters as positive and negative controls, fused to the GUS coding region, were used to coat microcarriers. The procedure for microcarrier preparation and precipitation of DNA onto microcarriers was essentially as described by Sanford et al., (1992).
Figure 5.7 Construction of the plasmid pZ2.47(-136)SATG in pBluescript SK. This plasmid was generated by replacing the intact -448/AflIII to +25/SphI region of pZ2.47SATG in pBluescript SK, by the mutated -448/AflIII to +25/SphI region from pZ2(-136)ES in pGem 1.0.
Transient-expression assay

Helium-driven PDS-1000/He particle gun was used to deliver tungsten particles coated with DNA into 3-d-old roots and 5-d-old leaves. Each petri plate containing root or leaf material was bombarded once with microcarriers coated with 5 \( \mu \)g of DNA. Conditions used were as in Held (1993). The gap distance between the rupture disk and macrocarrier was set to 8 mm, macrocarrier travel distance was set to 8 mm, and the target distance was set to 6 cm. Chamber vacuum was set at 28 inches Hg (0.06 atm), and the helium pressure was set at 1100 psi.

After bombardment the roots were incubated in darkness, and the leaves were incubated in light at 25\(^\circ\)C, for approximately 24 h. The leaves and roots were transferred to new petri dishes, using the Whatman paper they were placed on, and incubated in 5.0 ml of buffer [0.1 M Sodium phosphate (pH 6.8), 0.5 mM \( K_3Fe(CN)_6 \) and \( K_4Fe(CN)_6 \), 0.01 M EDTA, and 0.1% Triton-X-100] containing 0.5 mg/ml X-Glucuronidase (X-Gluc, Clontech), which is the histochemical substrate for the GUS enzyme. Roots and leaves in X-Gluc solution was incubated at 37\(^\circ\)C in the dark for 15-20 h. Blue spots were counted under a 10X stereo microscope.

Results

Using particle bombardment-mediated transient assay, Held (1993) had demonstrated that the first kbp of the \( zrp2 \) promoter most proximal to the transcription start site is capable of directing transcription in roots and leaf bases of maize.

To investigate whether cis-acting element/s located in the promoter region that showed differential pattern of DNA-protein interactions, with leaf and root nuclear protein
extracts, is responsible for root-preferential transcription of zrp2, this region was deleted and tested for transcriptional activity. Particle bombardment-mediated transient assays were used to measure the transcriptional activity of the 1 kbp and the 4.7 kbp promoter fragments, with or without the protein-binding region. The coding region of GUS was used as the reporter gene, and two other constructs 35S/GUS and a promoterless -35S/GUS were used as controls.

In bombardment experiments, the transiently expressing cells appear to correspond to those cells in which the particle has hit the nucleus (Yamashita et al., 1991). In this assay transcriptional activity of the promoters were quantified by counting the number of blue spots appeared on the leaves and roots, after incubation with the X-Gluc substrate, as done by Held, 1993. Quantification by counting the blue spots has proven to be reliable as the histochemical GUS enzyme assay in analysis of a rice promoter in stable transformants (Luan and Bogorad, 1992). Moreover, this method avoids error due to background fluorescence, probably contributed by contaminating epiphytic or endophytic bacteria or secondary compounds in plants (Jefferson, 1987; Held, 1993).

The transcriptional activity of the 1 kbp and the full 4.7 kbp zrp2 promoters were compared with corresponding mutated promoters in roots (Fig. 5.8), and leaves (Fig. 5.9). The mutated 1 kbp promoter with the protein-binding region deleted, showed five-fold-less activity than the 1 kbp promoter in roots (Fig 5.8). However the comparison of promoter activity between the 4.7 kbp and the mutated 4.7 kbp zrp2 promoter, did not show a significant difference (Fig. 5.8). In accordance with the observations made by Held (1993), the 35S promoter showed the highest transcriptional activity and the 4.7 kbp promoter showed approximately 3-fold-higher activity than the 1 kbp promoter fragment.
Figure 5.8 The transcriptional activity of 1 kbp and 4.7 kbp zrp2 promoters, with (+) or without (-) the protein-binding region, in roots. Four independent experiments were performed with the -35S, the 35S, the 1 kbp and the mutated 1 kbp promoter constructs, and two independent experiments were performed with the 4.7 kbp and the mutated 4.7 kbp promoter constructs, with each treatment replicated 2 to 4 times within an experiment. The activity of promoters are normalized to the activity of the 35S promoter. The mean percent of the promoter activity and the standard error of the mean are shown in the graph.
Figure 5.9 The transcriptional activity of the 1 kbp and 4.7 kbp zrp2 promoter, with (+) or without (-) the protein-binding region in leaf base and blade regions. Three independent experiments were performed with the -35S and the 35S promoter constructs and two independent experiments were performed with the intact and mutated 1 kbp and 4.7 kbp promoter constructs, with each treatment replicated 2 to 4 times within an experiment. The activity of promoters are normalized to the activity of the 35S promoter. The mean percent of the promoter activity and the standard error of the mean are shown in the graph.
The promoterless -35S construct showed virtually no transcriptional activity, confirming that the blue spots are a result of the transcriptional activity of the promoters.

In leaves, both 1 kbp and the 4.7 kbp promoters showed activity primarily in the basal region of the leaf, while the leaf blade area contained very few spots. Conversely, the 35S promoter activity appeared throughout the entire leaf. This pattern of distribution of the spots are similar to that observed by Held (1993). Because of this uneven distribution of the spots, the number of spots in leaf base and leaf blade areas were taken into account separately, and were represented as a percentage of the number of spots appeared on the corresponding region of leaves, bombarded with the 35S promoter. Comparison between the intact and the mutated promoters (both 1 kbp and 4.7 kbp), showed no significant difference in transcriptional activity in both leaf base and leaf blade areas (Fig. 5.9). The 4.7 kbp promoter was less than two-fold more active than the 1 kbp promoter in leaves. The activity of the 35S promoter remained the highest, and the promoterless -35S construct showed insignificant level of transcriptional activity as in roots (Fig. 5.9).

Discussion

Investigations of the zrp2 promoter for DNA-protein interactions using gel-shift assays, revealed that a 121 bp, -379 to -258 region (relative to the transcription start site), interacts differentially with nuclear proteins from leaves and roots (Chapter 3). Comparison of the transcriptional activity of the 1 kbp promoter fragment, with or without the above protein-binding region, using the particle bombardment-mediated transient assays in roots, showed five-fold-less activity when the protein-binding region was deleted (Fig. 5.8). These data indicate that this region is important in the level of transcription of
the *zrp2* gene in roots. Analysis of the sequence of this region revealed three 8 bp repeats and two 9 bp repeats, consisted entirely of A and T residues located in this region (Chapter 3). These results may indicate the presence of a enhancer element in this region, which is capable of enhancing the transcriptional level of the *zrp2* gene in roots.

Comparison of the transcriptional activity of the 4.7 kbp promoter, with or without the protein-binding region in roots, showed no significant difference in activity (Fig. 5.8). If this region contains an enhancer element for roots, deletion of this element should show reduced activity, unless there are other upstream sequences that are capable of enhancing the transcription in the absence of the proximal enhancer element. It is possible that the upstream sequences may contain redundant, long-range enhancer elements. Presence of long-range enhancer elements may explain, or may contribute to the higher activity of the 4.7 kbp promoter, compared to the 1 kbp promoter (Fig. 5.7 and Held, 1993). Analysis of the sequence of the upstream region revealed the presence of two additional copies of the 8 bp repeat, located at -1041 and -2053, relative to the transcription start site (Chapter 4). Together, these data suggest that this motif may function as an enhancer element in roots.

In leaves, the 1 kbp and the 4.7 kbp promoters showed no significant difference in the transcriptional activity when the protein-binding region was deleted. Absence of a significant decrease in transcriptional activity between the intact and the mutated 1 kbp promoter in leaves, suggests that the above putative enhancer element located in this region, functions in roots but not in leaves and therefore, is not a general transcription factor. Moreover, if this region contained a repressor element for leaves, deletion of this region would have resulted in increased transcriptional activity, unless there is a redundant negative element located inside the 1 kbp region. These results indicate that this region alone does not act as a negative element or as an enhancer element in leaves.
Comparison of figures 5.8 and 5.9 show that the overall transcriptional activity of the zrp2 promoters are higher in leaves than in roots. However, the transcriptional activity between root and leaves are not directly comparable, as they are represented as a percentage of the activity of the 35S promoter, which could be different between roots and leaves. Moreover, the differences in the shape, characteristics of the surfaces, surface area, and the hardness (degree of lignification) of roots and leaves contribute to differences in microparticle penetrating efficiency (Hunold et al., 1994).

In transient expression systems, the DNA need not integrate into the host genome. Therefore, the expression of the transferred gene will not depend on neighboring sequences in the recipient DNA, and thus there will be no position effects (Kuhlemeier et al., 1987). This is an advantage of transient expression systems over stable transformation. Conversely, the absence of chromatin structure around the introduced DNA may influence the regulation of transcription, thus deviating from its normal regulation of transcription in vivo. Moreover, the inability to control the copy number of the DNA introduced in particle bombardment-mediated transient assay system, is a disadvantage over stable transformants where a gene can be stable and present in single copy. Presence of multiple copies of promoter DNA in a single nucleus is not reflective of the in vivo condition. This could result in an excess of cis-elements in comparison to trans-acting protein factors in a nucleus, abolishing the delicate balance between the binding factors and binding sites which is important in regulation of transcription. The affect could be drastic on repressor elements which inhibit transcription when bound by repressor proteins. Excess of cis-acting repressor elements may titrate out the repressor proteins present in a nucleus, allowing the excess copies of the promoter to be transcribed without hindrance. This could be the situation with the zrp2 promoter in leaves (Held, 1993). This phenomenon does not
affect enhancer elements as much, because even if the activator proteins get titrated out, due to the promoter fragments that are already bound by the activator protein, final result would be an enhancement of transcription.

Moreover, an excess of cis-elements may prevent the competition for binding between specific and related protein factors. Therefore, the abundance or the binding affinity of the specific factor, which normally would play an important role in specificity of binding, may not be sufficient to prevent other factors from binding to the cis-element.
CHAPTER 6: GENERAL SUMMARY AND FUTURE GOALS

The main goal of this study is to identify regions of the \textit{zrp2} promoter, that are responsible for or associated with root-preferential transcription. Promoter DNA sequences from maize that could direct or enhance the transcription preferentially in roots may be useful in genetic engineering of agronomically important cereal grains.

Interactions between DNA and protein factors are known to modulate regulation of transcription. Therefore, to identify regions of the \textit{zrp2} promoter that positively regulate transcription preferentially in roots as compared to leaves, interactions between \textit{zrp2} promoter DNA and protein factors from crude nuclear protein extracts of leaves and roots were investigated to detect any DNA-protein binding activity that is different between roots and leaves. The first kbp of the \textit{zrp2} promoter region proximal to the transcription start site was chosen as it had been shown that this 1 kbp promoter fragment is capable of directing transcription in roots (Held, 1993), and generally elements that are responsible for organ or root-specific transcription are located within the first kbp of the promoter region.

As a control experiment for the presence of binding activity in the crude nuclear protein extracts of 3-d-old roots and 5-d-old leaves from maize, interactions between nuclear extracts and the \textit{ocs}-element of octapine synthase gene were analyzed, prior to the investigation of DNA-protein interactions with the \textit{zrp2} promoter (Chapter 1). Results of this investigation confirms the capability of factors in the crude nuclear protein extracts of maize roots and leaves to interact with its specific binding site.

Gel mobility-shift assays with the 1 kbp \textit{zrp2} promoter exhibited a differential DNA-protein binding pattern between root and leaf nuclear protein extracts, with the DNA sequences from -472 to -180, relative to the transcription start site (Chapter 3). The
protein-binding region was delimited to a 121 bp region (from -379 to -258) which showed 72% nucleotide identity with a root-preferentially expressed α-tubulin gene from maize. The high nucleotide identity in this region may indicate the possibility of this region's involvement in root-preferential transcription. Two DNA-protein complexes were observed with each nuclear protein extract. The complex with higher electrophoretic mobility (Z2BP1) was observed with both leaf and root nuclear protein extracts. The complexes Z2BP2 and Z2BP3 were observed exclusively with root and leaf nuclear protein extracts, respectively. Competition assays with non-labeled specific and non-specific competitor DNA showed that these interactions were specific. The use of a shorter fragment such as the 73 bp DNA fragment in competition assay could have exhibited the competition assay with root nuclear extracts more clearly, as the separation between the complexes and the free probe is better with the shorter probe. Further, use of multiple non-specific competitors could have further demonstrated the specificity of the DNA-protein interaction. Specifically, the use of an AT-rich non-specific competitor could have clearly shown whether the protein factors bind to the specific sequence elements or to any AT-rich DNA. However, absence of binding to the 36 bp or the 40 bp DNA fragments which are highly AT-rich shows that the interactions are likely to be sequence-specific.

The above results indicate that the promoter fragment interacts with nuclear factors in an organ-specific manner. These organ-specific differences in DNA-protein interactions correlate with the observed organ-preferential transcription activity (Held, 1993), thus providing circumstantial evidence for this promoter region's involvement in regulation of transcription in an organ-preferential manner. Two hypotheses that can be derived from these observations are: 1) the interactions between promoter DNA and protein factors that are responsible for the formation of the Z2BP2 complex enhance the level of transcription
in roots; and/or 2) the interactions responsible for the formation of the Z2BP3 complex represses the level of transcription in leaves. Furthermore, the high nucleotide identity of the protein-binding region with a region of another root-preferential gene provides additional support for this region's involvement in root-preferential transcription.

The effect of the protein-binding region on the level of transcription of the 1 kbp and the total 4.7 kbp promoters were analyzed by comparison of the transcriptional activity of promoter/GUS fusion constructs in particle bombardment-mediated transient assays in maize root and leaves. The transcriptional activity of the 1 kbp promoter, with the protein-binding region deleted, showed 5-fold-less activity than the intact 1 kbp promoter in roots, whereas no significant difference in activity was observed in leaves, indicating the presence of a positive element in this region, which acts as an enhancer, specifically in roots as compared to leaves.

Comparison of the transcriptional activity of the 4.7 kbp promoter, with or without the protein-binding region in roots showed no significant difference in activity, indicating the presence of other sequence elements in the upstream region, which act as long-range enhancers capable enhancing the transcription in the absence of the proximal putative enhancer element. Presence of such long-range enhancer(s) may explain the 3-fold-higher transcriptional activity of the 4.7 kbp promoter than the 1 kbp promoter. Presence of additional copies of the 8 bp repeat (TATTTTTTA) at -2053 and -1041 may suggest possible involvement of this motif as the putative enhancer element in roots.

A significant level of transcriptional activity was observed in roots and bases of leaves as compared to very low activity in leaf blades. This led to the analysis of DNA-protein interactions between leaf bases and leaf blades to investigate whether there are any differences in binding activity between leaf bases and blades and any similarity between
leaf bases and roots. The data indicate that Z2BP1 is formed in abundance with nuclear protein extracts from roots and leaf bases, and barely detectable with nuclear protein extracts from leaf blades. Consequently, the higher transcriptional activity in roots and leaf bases correlates with the presence of Z2BP1, and the absence of transcriptional activity in leaf blades correlate with the absence (or very low abundance) of Z2BP1, suggesting an involvement of Z2BP1 in transcriptional activity in transient assays.

The protein-binding region was subsequently delimited to a 73 bp region from -330 to -258, without a loss of binding activity. Analysis of the sequence of the 73 bp region showed the presence of two 9 bp direct repeats (5' AATAATTTA 3') and three 8 bp repeats (TATTTTTA) consisting only of A and T residues. Generally, both AT-rich regions and repeated sequences are known to interact with protein factors regulating transcription of eukaryotes.

Attempts to further delimit the binding sites resulted in loss of binding activity. Of the two double stranded oligonucleotides spanning the 73 bp fragment, only the 38 bp DNA fragment (3' end) showed binding activity. With the 38 bp fragment, only Z2BP1 and Z2BP3 were observed with Z2BP3 in low abundance. Decreased binding activity may be a result of loss of sequences that participate or facilitate binding, probably located in the periphery of the specific binding site. This observation lead to the speculation that the sequences responsible for the formation of Z2BP3 may be located near the 5' of the 38 bp DNA fragment. Inability to detect Z2BP2 with either 36 bp or 38 bp DNA fragments lead to the hypothesis that the binding site may have been divided by the separation of the 73 bp region into two fragments. However, the absence of Z2BP2 could be due to a requirement of sequences located in both fragments for binding. Some possible simplest models for
interactions between the zrp2 promoter and transcription factors were proposed based on these observations.

Conserved AT-rich sequence motifs are present in promoter regions of a number of plant genes. The 9 bp direct repeat ATTTAATAAA shows similarity to cis-elements having the consensus sequence T/AAT$_{3-4}$A$_{1-2}$. The 8 bp repeat TATTTTT closely resembles the consensus sequence of the conserved AT-1 box, AATATTTTTATT. AT-rich cis-elements have been reported to act as positive or negative regulatory elements. Therefore, a definite role for AT-rich sequence elements in transcription of plant genes remains unclear.

According to Forde (1994) there are two distinct types of nuclear proteins (HMGI-like and AT-rich binding factors-ATBFs) that interact with AT-rich upstream elements. Based on the criteria of Forde (1994), according to the characteristics observed, complexes Z2BP3 Z2BP2 and probably Z2BP1 fall in to the category of ATBFs. Binding factors of ATBF-type are widespread among dicots and evidence for occurrence is limited in monocots (Forde, 1994). Based on this, the nuclear factors responsible for the complexes Z2BP3, Z2BP2 and probably Z2BP1 are among the first few ATBF-type binding factors in monocots.

Further investigations of the effect of protein-binding sequences on transcription could be accomplished by gain-of-function studies where several copies of the protein-binding sequence are fused to the 5' end of a minimal promoter/reporter gene fusion constructs and transcriptional activity is assayed. The exact binding sites for these complexes could be investigated by footprint analysis. For this an in vitro method such as the DNase I footprint or an in vivo method such as DMS footprinting can be used. The exact nucleotide requirement for each binding site could be investigated by modifying the site with point mutations and performing gel mobility-shift assays. In parallel, these mutated DNA
fragments could be used in gain-of-function studies to determine exactly which DNA-protein interaction is responsible for root-preferential transcription. However, both gel mobility-shift assays and particle bombardment-mediated transient assays could deviate from the normal \textit{in vivo} situation, due to the introduction of excess of cis-elements. Therefore, use of stably transformed maize plants with \textit{zrp2} promoter fragments/reporter gene fusion constructs, would be desirable for analysis of transcriptional regulation.

There are other direct approaches to test whether these protein factors actually play a functional role in transcription. One way is to use an \textit{in vitro} transcription system and test the transcriptional activity in the presence or the absence of the purified transcription factor. To get purified transcription factors, the genes encoding the transcription factor must be cloned first and expressed in bacteria or yeast. The cloning is commonly done by screening a cDNA expression library from the particular organ of the plant with labeled DNA fragment containing the specific cis-element.

Another way of determining the effect of a transcription factor on transcription is to produce and investigate plants containing the antisense gene, of the gene encoding the transcription factor. As the antisense RNA form a duplex with the mRNA rendering it unstable, the final protein product will be drastically reduced. Therefore, the effect of very low levels or the absence of the transcription factor on transcription could be assayed.
LITERATURE CITED


My sincere gratitude goes to my major advisor, Dr. James Rothbart, for his invaluable advice, support, and encouragement during my tenure as a graduate student. My deep appreciation is also extended to my committee members, Dr. Clifford Cotterill, Dr. Jennifer Bodo, Dr. Dennis Larsen, and Dr. David Hanauer, for their insights and valuable time.

I would like to acknowledge several of the key people, past and present, for their contributions. My special thanks go to Dr. Linda Barnes, Dr. David Higgins, and Dr. Donna Perlin for being good friends. I also wish to thank Dr. Dennis Bollman, Victor Vong, Larsen, Michelle, Dr. Mike Baldwin, Dr. Kayla Gecky, and Dr. James Sibley.

My warmest gratitude goes to my parents who are very much responsible for my educational opportunities and for my brother and sister without whose help I would not have been possible.

Finally, I would like to express my deepest appreciation to my husband and devoted friend, Alphonse for his love, understanding and patience, and to my little girl, Raino, for being so sweet.
ACKNOWLEDGEMENTS

My sincere gratitude goes to my major advisor, Dr. James Colbert, for his invaluable advice, support and encouragement during the course of my graduate studies. My deep appreciation is also extended to my committee members, Dr. Clifford LaMotte, Dr. Martin Spalding, Dr. Drena Larson and Dr. David Hannapel for their insights and valuable time.

I would like to acknowledge members of our lab, past and present, for their contributions. My special thanks goes to Dr. Linda Barnes, Dr. David Higgs, and Dr. Bruce Held for being good friends. I also wish to thank Dr. Dennis Byrne, Victor Nunez, Lathika Moragoda, Dr. Iffat Rahim, Dr. Kevin Seeley, and Dr. Isaac John.

My unreserved gratitude goes to my parents who are very much responsible for my educational achievements, and for my brother and sisters without whom this would not have been possible.

Finally, I would like to extend my deepest appreciation to my husband and dearest friend, Anslem for his love, understanding and patience, and to our little girl, Roshanthi for being so content.
APPENDIX

Restriction map of the 2.6 kbp zrp2 promoter:

```
+-------------------+-------------------+-------------------+-------------------+-------------------+
| GGGCCCAAAGGAAGGAAGTACTTCAGTGATCAAGATGTGATGTTCTCCGATGGGTATGA | |
| CCGGGTTTTCTCTCTCTACTGGAAGTCCTAGTTCTACAACCTACAAGGGACTACCATA | |
| CAGCTAACCTGAGTAGGTGGGGTGAACTTATCTACTCTGTGAGTCTTAGGGATGAAGAGT | |
| GTCGATTGGACTCATCCACCCCACTTGAATAGATGAGACACTCAGAATCCCTACTTCTCA | |
+-------------------+-------------------+-------------------+-------------------+-------------------+
```
ATGATATATGTGATGTATATATGTGATGTATATGTGACATATGTGATGTATATGTGGTAT
---+-----+---+---+---+---+---+---+---+---+---+---+---+---+---+---
TACTATATACACTACATATATACACTACACTACATACACTACATACACTACATACATACACCATA

CTTTTGTTTGGATGGAATAGAGAAAGCAATAAAATGTGTAATCTGCTGTCATCTTTT
---+-----+---+---+---+---+---+---+---+---+---+---+---+---+---+---
GAAAACAAACAACCTACCTTATCTCTTTTCCTTTATTTTTACACATATGACCGAGTGAAAC

TCGAGTGTAACACTCGGCAAAAAGGTGCTTTGCCGAGTGTTAGGGCCATAGCATCGGTA
---+-----+---+---+---+---+---+---+---+---+---+---+---+---+---+---
AGCTCACAATCGAGCCGTTTTCACGAACCGCGCTCAACATCGGCTGAGCGCAT

GAGAACCAATACTTAGGCACCGGTAAAGCTTTTTTGGCGAGTGTTGTGGCCCTGGCACTC
---+-----+---+---+---+---+---+---+---+---+---+---+---+---+---+---
CTCTTGGTTATGAAATCCGTGGCCATCGAAAAACCCGCTCACAACACCGGGACCGTGAGG

MT as ep N I 4 d I 5 e II I

B s t M T a s
B F A1 e Bp
c c0 I s4
c c7 I r5
I II I II

M a S H a Ca
t e u ve
t a I 9 I
q I 6 JI
I I I II

H i S EH aCBcaS
D N BB i n C
r D BlAssM d Av M uvsoec D

d d a agars I li w 9iaRir d
I e n LeWFp I uJ o 6JIIF e
I I I VIII I II I IIIII I

GAGAACCAATACTTAGCGACCCTAAAGCCTTTTTTGGCGAGTGTGTTGGCCCTGGC ACTC
---+-----+---+---+---+---+---+---+---+---+---+---+---+---+---+---
CTCTTGGTTATGAAATCCGTGGCCATTTCCGAAAAACCCTCAACACCCGG ACCCTGAG
CGAGGATTGAAACGGGGTAATTTGGTCCCCATGTGCCCGTCCCGCGAACTTCTCTTGATC
GCTCCTAACTTGTCCCATTAACCAAGGGGTACACGGGCAGGGCGCTTGAAGAGAACTAG
TAAATTAGTCTATTTCCATGTAAAAACTATACTAAAAATTTAATACACAGTCTATTTATAA
ATTTAATCAGATAAAGGTACAATTTTGATATGATTTTTAAATTATGTGTCAGATAATATT
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
TTATCGTTTGATTTAAGATTTCAACTGTAACATGAAACTAAATTTAGACCAAACAAGTTC
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
TTATCGTTTGATTTAAGATTTCAACTGTAACATGAAACTAAATTTAGACCAAACAAGTTC
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
TTATCGTTTGATTTAAGATTTCAACTGTAACATGAAACTAAATTTAGACCAAACAAGTTC
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
TTATCGTTTGATTTAAGATTTCAACTGTAACATGAAACTAAATTTAGACCAAACAAGTTC
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
TTATCGTTTGATTTAAGATTTCAACTGTAACATGAAACTAAATTTAGACCAAACAAGTTC
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
TTATCGTTTGATTTAAGATTTCAACTGTAACATGAAACTAAATTTAGACCAAACAAGTTC
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
TTATCGTTTGATTTAAGATTTCAACTGTAACATGAAACTAAATTTAGACCAAACAAGTTC
TTATATTCATTTGATATAATAATTTGAATTTGACTCTTAATATCGTATTTTTTCCTAAC
AATAGCAAACTAAATTCATAAAGTTGATGCACTTTGTAATTAAAAATCCTGTGTTGTTCAAG
<table>
<thead>
<tr>
<th>N</th>
<th>H</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>B 1RK</td>
<td>iIT</td>
<td>a</td>
</tr>
<tr>
<td>a asp f</td>
<td>nf</td>
<td>e</td>
</tr>
<tr>
<td>n Ian c</td>
<td>fi</td>
<td>I</td>
</tr>
<tr>
<td>I VII I</td>
<td>II</td>
<td>I</td>
</tr>
</tbody>
</table>

```
GTAATGAGATATAGAATAAGGTACCCTATAGAATAAAAGAATCTTTCTCTGAAAAGTCTGA
-173
```

```
CATTTCTATATCTTTATCCATGGGATATCTTTATTTTTCTTAGAAAGAGACTTTTTCAGACT
```

<table>
<thead>
<tr>
<th>E</th>
<th>co</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBS</td>
<td>4</td>
</tr>
<tr>
<td>S Rasn</td>
<td>7Hafs</td>
</tr>
<tr>
<td>u seaa</td>
<td>Heap</td>
</tr>
<tr>
<td>n aIAB</td>
<td>IaING</td>
</tr>
<tr>
<td>I IIIII</td>
<td>IIIII</td>
</tr>
</tbody>
</table>

```
CATTACTCTATATCTTTATCCATGGGATATCTTTATTTTTCTTAGAAAGAGACTTTTTCAGACT
-113
```

```
GCACTGATTATATTATATTATATTTTTCTTCGCGACCTGACCTCTACGAG
```

<table>
<thead>
<tr>
<th>B</th>
<th>T</th>
<th>C</th>
<th>B</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAI B</td>
<td>s</td>
<td>a</td>
<td>v</td>
<td>s</td>
</tr>
<tr>
<td>sct p</td>
<td>r</td>
<td>q</td>
<td>i</td>
<td>a</td>
</tr>
<tr>
<td>lia m</td>
<td>D</td>
<td>I</td>
<td>J</td>
<td>J</td>
</tr>
<tr>
<td>III I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
</tbody>
</table>

```
CGTACGTAAATAAGATATATAATAAAAAATCTACAAGAAGACTGGGACTGGAGATGCTC
-113
```

```
GCATGCATTATCTATATTATTTTTATTTTTATTTTTCATCTCGGACTGCACCTCTACGAG
```

<table>
<thead>
<tr>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

```
GTATATGCAGAATGCCCAGTGTCTATGATAATAGCACCACCTCGGGCTGCGGAGACATGAAACGG
-53
```

```
GATATACGCCCTTACGGACACGGAATATTATCGGTGGAAGCAGCAGAGGCTTCTGTACTTGCC
+1
```

<table>
<thead>
<tr>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

```
E     Hcv a aNS
C     geic Isp
i     afR8 Iph
I     IIIII III
```

```
CGGACGCAGTGTGCATGC
+9
```

```
GCCTGCGTCACACGTCAG
```
### Enzymes that do cut:

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>AccI</td>
<td>AcII</td>
<td>AflIII</td>
<td>AgeI</td>
<td>AluI</td>
<td>AlwI</td>
<td>Alw2I</td>
<td>ApaI</td>
<td>ApoI</td>
<td>AvaI</td>
</tr>
<tr>
<td>ApoI</td>
<td>AvaI</td>
<td>AveII</td>
<td>BaeI</td>
<td>BanI</td>
<td>BanII</td>
<td>BbsI</td>
<td>BbvI</td>
<td>BccI</td>
<td>Bce8III</td>
</tr>
<tr>
<td>Bce8III</td>
<td>Bcefl</td>
<td>BcgI</td>
<td>BfaI</td>
<td>BpmI</td>
<td>Bpu10I</td>
<td>BsaI</td>
<td>BsaAI</td>
<td>BsaAI</td>
<td>BsaWI</td>
</tr>
<tr>
<td>BsaAI</td>
<td>BsaAI</td>
<td>BscGI</td>
<td>BsgI</td>
<td>BsiI</td>
<td>BsiEI</td>
<td>BslI</td>
<td>BsmAI</td>
<td>BsmAI</td>
<td>BsmFII</td>
</tr>
<tr>
<td>BsmAI</td>
<td>BsmFII</td>
<td>Bsp24I</td>
<td>Bsp1286I</td>
<td>BspGI</td>
<td>BspLUIII</td>
<td>BspMI</td>
<td>BsrI</td>
<td>BsrI</td>
<td>BsrDII</td>
</tr>
<tr>
<td>BsrDII</td>
<td>BsrDII</td>
<td>Bst1107I</td>
<td>BstYI</td>
<td>Cac8I</td>
<td>Clal</td>
<td>CviJI</td>
<td>CviRI</td>
<td>DdeI</td>
<td>DpnI</td>
</tr>
<tr>
<td>DdeI</td>
<td>DpnI</td>
<td>DraI</td>
<td>DraiII</td>
<td>DsaI</td>
<td>EaeI</td>
<td>Eam1105I</td>
<td>EarI</td>
<td>EarI</td>
<td>Eco47III</td>
</tr>
<tr>
<td>EarI</td>
<td>Eco47III</td>
<td>Eco57I</td>
<td>EcoO109I</td>
<td>EcoRI</td>
<td>EcoRII</td>
<td>EcoRV</td>
<td>FauI</td>
<td>FokI</td>
<td>GdiII</td>
</tr>
<tr>
<td>FauI</td>
<td>FokI</td>
<td>GdiII</td>
<td>HaeII</td>
<td>HaeIII</td>
<td>HgaI</td>
<td>HgiEII</td>
<td>HhaI</td>
<td>HincII</td>
<td>HindIII</td>
</tr>
<tr>
<td>HincII</td>
<td>HindIII</td>
<td>HinfI</td>
<td>HphI</td>
<td>Itai</td>
<td>KpnI</td>
<td>MaeII</td>
<td>MaeIII</td>
<td>MboII</td>
<td>MmeI</td>
</tr>
<tr>
<td>MboII</td>
<td>MmeI</td>
<td>MnlI</td>
<td>MseI</td>
<td>MsII</td>
<td>MspI</td>
<td>MspAI</td>
<td>MwoI</td>
<td>NciI</td>
<td>NdeI</td>
</tr>
<tr>
<td>NciI</td>
<td>NdeI</td>
<td>NheI</td>
<td>NlaIII</td>
<td>NlaIV</td>
<td>NsiI</td>
<td>NspI</td>
<td>NspV</td>
<td>PleI</td>
<td>Psp5II</td>
</tr>
<tr>
<td>PleI</td>
<td>Psp5II</td>
<td>RcaI</td>
<td>RleAI</td>
<td>RsaI</td>
<td>SacII</td>
<td>SalI</td>
<td>Sau96I</td>
<td>Sau3AI</td>
<td>ScaI</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>ScaI</td>
<td>ScrFI</td>
<td>SfαNI</td>
<td>SfcI</td>
<td>SnaBI</td>
<td>Sphi</td>
<td>SpI</td>
<td>SunI</td>
<td>TaqI</td>
</tr>
<tr>
<td>SunI</td>
<td>TaqI</td>
<td>TaqII</td>
<td>TfiI</td>
<td>Thai</td>
<td>Tsp45I</td>
<td>Tsp509I</td>
<td>Tth111II</td>
<td>VspI</td>
<td>XcmI</td>
</tr>
<tr>
<td>VspI</td>
<td>XcmI</td>
<td>XmnI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Enzymes that do not cut:

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>AatII</td>
<td>AflII</td>
<td>Alw44I</td>
<td>AlwNI</td>
<td>ApaBI</td>
<td>AscI</td>
<td>AvrII</td>
<td>BanHI</td>
</tr>
<tr>
<td>BclI</td>
<td>BglI</td>
<td>BglIII</td>
<td>Bpu1102I</td>
<td>BsaBI</td>
<td>BsaHI</td>
<td>BseRI</td>
<td>BsmI</td>
</tr>
<tr>
<td>BspEI</td>
<td>BsrBI</td>
<td>BsrGI</td>
<td>BssHII</td>
<td>BstEII</td>
<td>BstXI</td>
<td>Bsu36I</td>
<td>DraIII</td>
</tr>
<tr>
<td>DrdI</td>
<td>EcoNI</td>
<td>Esp3I</td>
<td>FseI</td>
<td>FspI</td>
<td>HaeI</td>
<td>HpaI</td>
<td>MluI</td>
</tr>
<tr>
<td>MscI</td>
<td>MunI</td>
<td>NarI</td>
<td>NcoI</td>
<td>NgoAIV</td>
<td>NotI</td>
<td>NruI</td>
<td>PacI</td>
</tr>
<tr>
<td>Pf1MI</td>
<td>Pf1MI</td>
<td>PmeI</td>
<td>PmlI</td>
<td>PshAI</td>
<td>Psp1406I</td>
<td>PstI</td>
<td>PvuI</td>
</tr>
<tr>
<td>PvuII</td>
<td>RsrII</td>
<td>SacI</td>
<td>SapI</td>
<td>SexAI</td>
<td>SflI</td>
<td>Sgfl</td>
<td>SgrAI</td>
</tr>
<tr>
<td>Smal</td>
<td>SpeI</td>
<td>SrfI</td>
<td>Sse8387I</td>
<td>StuI</td>
<td>StyI</td>
<td>Swai</td>
<td>Tth111I</td>
</tr>
</tbody>
</table>