ISOLATION, IDENTIFICATION AND CHARACTERIZATION OF

*Os*RGLP2 PROMOTER BINDING PROTEIN(S)

FARAH DEeba
06-arid-773

Department of Biochemistry
Faculty of Sciences
Pir Mehr Ali Shah
Arid Agriculture University Rawalpindi
Pakistan
2016
ISOLATION, IDENTIFICATION AND CHARACTERIZATION OF

*OsRGLP2* PROMOTER BINDING PROTEIN(S)

by

FARAH DEEBA
(06-arid-773)

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biochemistry

Department of Biochemistry
Faculty of Sciences
Pir Mehr Ali Shah
Arid Agriculture University Rawalpindi
Pakistan
2016
CERTIFICATION

I hereby undertake that this research is an original and no part of this thesis falls under plagiarism. If found otherwise, at any stage, I will be responsible for the consequences.

Name: Farah Deeba  Signature:______________
Registration No.: 06-arid-773  Date: ________________

Certified that the contents and form of thesis entitled “Isolation, Identification and Characterization of OsRGLP2 Promoter Binding Protein(s)” submitted by Ms. Farah Deeba have been found satisfactory for the requirement of the degree.

Supervisor:______________ (Prof. Dr. S. M. Saqlan Naqvi)
Co-supervisor:______________ (Dr. Tariq Mahmood)
Member:______________ (Dr. Ghazala Kaukab)
Member:______________ (Prof. Dr. Tariq Masud)

Chairman:________________________
Dean:________________________

Director Advanced Studies:________________________
DEDICATED TO MY PARENTS, SIBLINGS AND UMAR

FOR THEIR ENDLESS LOVE, SUPPORT AND ENCOURAGEMENT
# CONTENTS

| LIST OF FIGURES                  | xi  |
| LIST OF PLATES                  | xii |
| LIST OF TABLES                  | xiv |
| LIST OF ABBREVIATIONS           | xv  |
| ACKNOWLEDGEMENTS                | xvii|
| ABSTRACT                        | xix |
| 1. INTRODUCTION                 | 1   |
| 2. REVIEW OF LITERATURE         | 7   |
| 2.1 GERMIN AND GERMIN LIKE PROTEINS (GLPs): THE DIVERSE MEMBERS OF CUPIN SUPERFAMILY | 7   |
| 2.2 BIOCHEMICAL NATURE OF GLPs  | 7   |
| 2.2.1 Oxalate Oxidase Activity  | 8   |
| 2.2.2 Superoxide Dismutase Activity | 8   |
| 2.2.3 ADP glucose Pyrophosphatase or Phosphodiesterase (AGPPase) Activity | 8   |
| 2.2.4 Polyphenol Oxidase Activity | 8   |
| 2.3 REGULATION OF GLPs EXPRESSION DURING VARIOUS STRESSES | 9   |
| 2.3.1 Role of GLPs in Biotic Stresses | 9   |
| 2.3.2 Role of GLPs in Abiotic Stresses | 9   |
| 2.4 GLPs PROMOTER ANALYSIS      | 10  |
| 2.5 CIS-ACTING REGULATORY ELEMENTS FOR ABIOTIC AND BIOTIC STRESSES | 11  |
| 2.5.1 Abscisic Acid Responsive Elements (ABRE) Motifs | 12  |
| 2.5.2 Drought Responsive Elements(DREs) Motifs | 12  |
2.5.3 Ethylene-Responsive Elements (EREs) Motifs
2.5.4 Anaerobic-Responsive Elements (AREs) Motifs
2.5.5 W-Boxes
2.6 TRANSCRIPTION FACTORS ARE INVOLVED IN BIOTIC AND ABIOTIC STRESSES
2.6.1 WRKY Transcription Factors
2.6.1.1 WRKY domain and classification
2.6.1.2 WRKY transcription factors in stresses
2.6.2 MYB Transcription Factors in Plants
2.6.2.1 DNA-binding specificity of MYB Factors
2.6.2.2 Role of MYB in abiotic stresses
2.6.3 DNA With One finger (DOF) Transcription Factors
2.6.3.1 Role of DOF in stresses
2.6.4 NAC (NAM/ATAF1/CUC2) Transcription Factors
2.7 STRATEGIES TO UNRAVEL THE FUNCTION OF PLANT PROMOTERS
2.7.1 Filter Binding Assays
2.7.2 DNase I Footprinting
2.7.3 Electrophoretic Mobility Shift Assays (EMSAs)
3 MATERIALS AND METHODS
3.1 SEQUENCE ANALYSIS OF OsRGLP2 PROMOTER
3.2 SEQUENCE RETRIEVAL AND IN SILICO ANALYSIS OF PROTEINS THAT PUTATIVELY BIND TO OsRGLP2 PROMOTER
3.3 CLONING OF PROTEINS THAT BIND TO OsRGLP2 PROMOTER
3.3.1 Primer Designing
3.3.2 Plant Material

3.3.3 RNA Isolation and Synthesis of cDNA

3.3.4 PCR Amplification of OsWRKY71 and OsMYB1 DNA Binding Domains

3.3.5 PCR Amplification of OsDOF18

3.3.6 Cloning of OsWRKY71, OsDOF18 and OsMYB1 DNA Binding Domain Regions

3.3.6.1 Transformation of ligation reaction in E. coli

3.3.6.2 Confirmation of the Positive Clones

3.4 OVER-EXPRESSION OF GST, GST-OsWRKY71 (167-271), GST-OsDOF18 (1-277), GST-OsDOF18 (1-148) and GST-OsMYB1 (88-222)

3.4.1 Transformation of Recombinant vectors in Expression Cells of E. coli

3.4.2 Expression of GST, GST-OsWRKY71, GST-OsDOF18 (1-277), GST-OsDOF18 (1-148) and GST-OsMYB1

3.5 PURIFICATION OF GST, GST-OsWRKY71 (167-271), GST-OsDOF18 (1-148) and GST-OsMYB1 (88-222)

3.5.1 Cell Lysis

3.5.2 Glutathione Sepharose Column Chromatography

3.5.3 SDS-PAGE

3.5.4 Cationic Exchange Chromatography

3.5.5 Purified GST, GST-OsWRKY71 (167-271), GST-OsMYB1 (88-222), and GST-OsDOF18 (1-148) Dialysis and Concentration Determination

3.6 PURIFIED PROTEINS CONFIRMATION BY MASS SPECTROMETRY

3.6.1 On Column Cleavage of GST Tagged Proteins

3.6.2 Analysis on Mass Spectrometer
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6.2.1</td>
<td>Target cleaning and matrix preparation</td>
<td>35</td>
</tr>
<tr>
<td>3.6.2.2</td>
<td>MALDI TOF Mass Spectrometry</td>
<td>35</td>
</tr>
<tr>
<td>3.7</td>
<td>ELECTROPHORETIC MOBILITY SHIFT ASSAYS (EMSAs) WITH PURIFIED RECOMBINANT PROTEINS</td>
<td>36</td>
</tr>
<tr>
<td>3.7.1</td>
<td>Probe Designing</td>
<td>36</td>
</tr>
<tr>
<td>3.7.2</td>
<td>Probes Labelling and Purification</td>
<td>37</td>
</tr>
<tr>
<td>3.7.3</td>
<td>Labelling Efficiency Determination</td>
<td>37</td>
</tr>
<tr>
<td>3.7.4</td>
<td>Binding Reaction Preparation and Optimization</td>
<td>39</td>
</tr>
<tr>
<td>3.7.5</td>
<td>Native Polyacrylamide Gel Electrophoresis (Native PAGE)</td>
<td>41</td>
</tr>
<tr>
<td>3.7.6</td>
<td>Blotting and Crosslinking</td>
<td>41</td>
</tr>
<tr>
<td>3.7.7</td>
<td>Chemiluminescent Signal Detection and X-ray Exposure</td>
<td>42</td>
</tr>
<tr>
<td>3.8</td>
<td>EMSA WITH NUCLEAR PROTEINS EXTRACT</td>
<td>43</td>
</tr>
<tr>
<td>3.8.1</td>
<td>Nuclear Proteins Extraction</td>
<td>43</td>
</tr>
<tr>
<td>3.8.2</td>
<td>Binding Reaction and Detection</td>
<td>44</td>
</tr>
<tr>
<td>3.9</td>
<td>EXPRESSION PROFILE OF TARGET GENES</td>
<td>45</td>
</tr>
<tr>
<td>3.9.1</td>
<td>Microarray Data Retrieval from GENEVESTIGATOR</td>
<td>45</td>
</tr>
<tr>
<td>3.9.2</td>
<td>Quantitative Real-Time PCR (QRT-PCR)</td>
<td>45</td>
</tr>
<tr>
<td>3.9.2.1</td>
<td>Primer designing</td>
<td>45</td>
</tr>
<tr>
<td>3.9.2.2</td>
<td>Plant growth conditions and stress treatments</td>
<td>47</td>
</tr>
<tr>
<td>3.9.2.3</td>
<td>RNA Extraction and determination of RNA quality</td>
<td>48</td>
</tr>
<tr>
<td>3.9.2.4</td>
<td>Real time PCR</td>
<td>48</td>
</tr>
<tr>
<td>3.9.2.5</td>
<td>Data analysis of qRT-PCR</td>
<td>49</td>
</tr>
<tr>
<td>3.10</td>
<td>ANALYSIS OF <em>E. COLI</em> TRANSFORMANTS IN ABIOTIC STRESSES</td>
<td>49</td>
</tr>
<tr>
<td>3.11</td>
<td>IN SILICO CHARACTERIZATION OF IDENTIFIED PROTEINS</td>
<td>51</td>
</tr>
</tbody>
</table>
## RESULTS AND DISCUSSION

### 4.1 IN SILICO ANALYSIS OF OsRGLP2 PROMOTER FOR STRESS RESPONSIVE REGULATORY MOTIFS AND CORRESPONDING BINDING PROTEINS

### 4.2 RETRIEVAL AND SEQUENCE ANALYSIS OF OsWRKY71, OsDOF18 AND OsMYB1 TRANSCRIPTION FACTORS

#### 4.2.1 Sequence Analysis of OsWRKY71

#### 4.2.2 Identification and Sequence Analysis of OsMYB1

#### 4.2.3 Identification and Sequence Analysis of OsDOF18

### 4.3 CLONING, OVEREXPRESSION AND PURIFICATION OF OsWRKY71, OsDOF18 AND OsMYB1 DNA BINDING DOMAINS

#### 4.3.1 Cloning of DNA Binding Domain

#### 4.3.1.1 Cloning of OsWRKY71 DNA binding domain

#### 4.3.1.2 Cloning of OsMYB1 DNA binding domain

#### 4.3.1.3 Cloning of OsDOF18 DNA binding domain

#### 4.3.2 Expression Level of Recombinant GST-OsWRKY71 (167-271) in Different E. coli Strains

#### 4.3.3 Purification of GST-OsWRKY71 (167-271) from BL21

#### 4.3.4 Expression Level of Recombinant GST-OsDOF18 (1-148) in Different E. coli Strains

#### 4.3.5 Purification of GST-OsDOF18 (1-148) from BL21

#### 4.3.6 Expression Level of Recombinant GST-OsMYB1(88-222) protein in different E. coli strains

#### 4.3.7 Purification of GST-OsMYB1 (88-222)

### 4.4 MALDI-TOF MASS SPECTROMETRY (MS) OF CLEAVED PROTEINS FRAGMENTS

### 4.5 BINDING OF RECOMBINANT PROTEINS WITH...
4.5.1 Labelling Efficiency

4.5.2 Analysis of Interaction of the GST-\(Os\)WRKY71 (167-271) with Putative WRKY Binding Motifs on \(Os\)RGLP2 Promoter

4.5.2.1 Probe containing one true and one core W-Box

4.5.2.2 Probe containing one true W-Box

4.5.2.3 Probe containing one core W-Box TGAC on both strands

4.5.3 Analysis of Interaction of the GST-\(Os\)MYB1 (88-222) with Putative MYB Binding Motifs on \(Os\)RGLP2 Promoter

4.5.3.1 Probe containing TAACCA motif

4.5.3.2 Probe containing AAACCA motif

4.5.4 Analysis of Interaction of the GST-\(Os\)DOF18 (1-148) with Putative DOF Binding Motifs on \(Os\)RGLP2 Promoter

4.5.4.1 Probe containing two AAAG motifs

4.5.4.2 Probe containing one AAAG motif

4.6 EXPRESSION ANALYSIS OF \(Os\)WRKY71, \(Os\)MYB1 AND \(Os\)DOF18 BY REAL TIME PCR

4.6.1 \(Os\)WRKY71 Gene Expression in Response to Various Abiotic Stresses

4.6.2 \(Os\)DOF18 Gene Expression in Response to Various Abiotic Stresses

4.6.3 \(Os\)MYB1 Gene Expression in Response to Various Abiotic Stresses

4.7 EXPRESSION PATTERNS Of GST-\(Os\)WRKY71, GST\(Os\)DOF18 AND GST-\(Os\)MYB1 IN \textit{E. coli} UNDER DIFFERENT ABIOTIC STRESSES

4.7.1 Expression Pattern of GST-\(Os\)WRKY71 in \textit{E. coli}

4.7.2 Expression Pattern of GST-\(Os\)DOF18 in \textit{E. coli}

4.7.3 Expression Pattern of GST-\(Os\)MYB1 in \textit{E. coli}
4.8 IN SILICO CHARACTERIZATION OF OsWRKY71, OsDOF18 AND OsMYB1

4.8.1 Intrinsic Disorder in OsWRKY71, OsDOF18 and OsMYB1  104

4.8.2 Protein-protein Interaction Network Analysis  106

SUMMARY  113

LITERATURE CITED  116

APPENDICES  140
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Position of different stress responsive cis-elements on 1107 bp OsRGLP2 promoter region</td>
<td>54</td>
</tr>
<tr>
<td>2</td>
<td>Protein sequence alignment of OsWRKY71 structure of DNA binding domain</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>Protein sequence alignment of OsMYB1 with different MYB proteins and structure of DNA binding domain</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>Protein sequence alignment of OsDOF18 with DOF proteins from different plants and structure of DNA binding domain</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Expression vector pGEX-4T1 for OsWRKY71, OsMYB1 and OsDOF18</td>
<td>62</td>
</tr>
<tr>
<td>6</td>
<td>MALDI-TOF spectrum of the cleaved OsWRKY71 (167-271)</td>
<td>75</td>
</tr>
<tr>
<td>7</td>
<td>MALDI-TOF spectra of the cleaved OsMYB1 fragments</td>
<td>76</td>
</tr>
<tr>
<td>8</td>
<td>Expression analysis of OsWRKY71</td>
<td>94</td>
</tr>
<tr>
<td>9</td>
<td>Expression analysis of OsDOF18</td>
<td>95</td>
</tr>
<tr>
<td>10</td>
<td>Expression analysis of OsMYB1</td>
<td>97</td>
</tr>
<tr>
<td>11</td>
<td>Intrinsic disorder (ID) prediction for OsWRKY71</td>
<td>105</td>
</tr>
<tr>
<td>12</td>
<td>Intrinsic disorder (ID) prediction for OsDOF18</td>
<td>107</td>
</tr>
<tr>
<td>13</td>
<td>Intrinsic disorder (ID) prediction for OsMYB1</td>
<td>108</td>
</tr>
<tr>
<td>14</td>
<td>Protein-protein interaction network of OsWRKY71 analyzed by String</td>
<td>110</td>
</tr>
<tr>
<td>15</td>
<td>Protein-protein interaction network of OsDOF18 analyzed by String</td>
<td>111</td>
</tr>
<tr>
<td>16</td>
<td>Protein-protein interaction network of OsMYB1 analyzed by String</td>
<td>112</td>
</tr>
</tbody>
</table>
# LIST OF PLATES

<table>
<thead>
<tr>
<th>Plate No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RNA isolated from 10 day old seedlings</td>
<td>62</td>
</tr>
<tr>
<td>2</td>
<td>Cloning of OsWRKY71 DNA binding domain</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>Cloning of OsMYB1 DNA binding domain</td>
<td>65</td>
</tr>
<tr>
<td>4</td>
<td>Cloning of OsDOF18 DNA binding domain</td>
<td>67</td>
</tr>
<tr>
<td>5</td>
<td>Expression and purification of GST-OsWRKY71 (167-271)</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>Expression and purification of GST-OsDOF18 (1-148)</td>
<td>71</td>
</tr>
<tr>
<td>7</td>
<td>Expression and purification of GST-OsMYB1 (88-222).</td>
<td>73</td>
</tr>
<tr>
<td>8</td>
<td>Labelling efficiency of DIG labeled oligos using a dot blot</td>
<td>78</td>
</tr>
<tr>
<td>9</td>
<td><strong>In vitro</strong> DNA binding ability of OsWRKY71 with OsRGLP2 fragment containing one core W-box and one true W box</td>
<td>81</td>
</tr>
<tr>
<td>10</td>
<td><strong>In vitro</strong> DNA binding ability of OsWRKY71 with OsRGLP2 fragment containing one true W box</td>
<td>82</td>
</tr>
<tr>
<td>11</td>
<td><strong>In vitro</strong> DNA binding ability of OsWRKY71 with OsRGLP2 fragment containing one core W-box on both strands</td>
<td>84</td>
</tr>
<tr>
<td>12</td>
<td><strong>In vitro</strong> DNA binding ability of OsMYB1 with OsRGLP2 fragment containing one MYB binding site TAACCA</td>
<td>87</td>
</tr>
<tr>
<td>13</td>
<td><strong>In vitro</strong> DNA binding ability of OsMYB1 with OsRGLP2 fragment containing one MYB binding site AAACCA</td>
<td>88</td>
</tr>
<tr>
<td>14</td>
<td><strong>In vitro</strong> DNA binding ability of OsDOF18 with OsRGLP2 fragment containing two AAAG boxes</td>
<td>90</td>
</tr>
<tr>
<td>15</td>
<td><strong>In vitro</strong> DNA binding ability of OsDOF18 with OsRGLP2 fragment containing one AAAG motif</td>
<td>92</td>
</tr>
<tr>
<td>Page</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Spot assay of BL/pGEX4T-1 and BL/OsWRKY71 recombinants</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Spot assay of BL/pGEX4T-1 and BL/OsDOF18 recombinants</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Spot assay of BL/pGEX4T-1 and BL/OsMYB1 recombinants</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Primer sequences used for cloning of OsWRKY71, OsMYB1 and OsDOF18 DNA binding domains and full length OsDOF18</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>OsRGLP2 promoter fragments used as probe in Electrophoretic mobility shift assays</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>Buffers screened for EMSA</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>Primers designed for Real Time PCR</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>Stress responsive cis-acting elements and their corresponding binding factors located on OsRGLP2 promoter and their role in different stresses</td>
<td>53</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

µL       Micro litre
ABREs    Abscisic Acid Responsive Elements
AREs     Anaerobic-Responsive Elements
BLAST    Basic Local Alignment Search Tool
bp       Base Pair
bZIP     basic leucine zipper
CSPD     Disodium 3-(4-methoxyspiro[1,2-dioxetane-3,2'-(5'-chloro)-tricyclo[3.3.1.13,7]decan]-4-yl)phenyl phosphate
DIG      Digoxigenin
DNA      Deoxyribo Nucleic Acid
dNTPs    Deoxyribo Nucleoside Triphosphate
DOF      DNA with One finger
DREs     Drought Responsive Elements
DTT      Dithiothreitol
EDTA     Ethylene diamine tetra acetic acid
EMSAs    Electrophoretic Mobility Shift Assays
G        Gram
GLPs     Germin like proteins
GST      Glutathione S-transferase
IPTG     IsoPropyl-beta-D-ThioGlactopyranoside
kb       Kilo base pair
L        Litre
LB       Lauria-Bartani
LiCl     Lithium Chloride
M        Molar
mA       Milli ampere
MCS      Multiple Cloning Sites
mg       Milli gram
MgCl₂    Magnesium chloride
mM       Milli molar
MYB  Myeloblastosis viral oncogene homolog
NAC  (NAM/ATAF1/CUC2)
NaCl  Sodium Chloride
ng  Nano gram
OD  Optical density
PCR  Polymerase Chain Reaction
pM  Pico molar
RNA  Ribonucleic Acid
RNAse A  Ribonuclease A
SDS  Sodium Dodecyl Sulphate
SSC  saline-sodium citrate
TE  Tris-EDTA
UV  Ultra violet
V  Voltage
ACKNOWLEDGEMENTS

All praise and thanks are due to ALMIGHTY ALLAH, WHO in His Infinite Mercy and Grace enabled me to complete the present thesis. I bow my head with all submission and humility by way of gratitude due to Almighty Allah. Glory and praise to our Last PROPHETHAZRAT MUHAMMAD (Peace be upon Him), Who is forever a symbol of direction and knowledge for whole humanity.

First of all, I would like to express my deep gratitude to my supervisor Prof. Dr. S. M. SaqlanNaqvi, Dean, Faculty of Sciences, for providing the visionary framework upon which this work was built and constantly challenging me to take this body of work to its maximal potential. I thank him for his enormous support and guidance and for believing in me and for his constant encouragement and advices.

I am very thankful to Higher Education Commission Pakistan for providing financial assistance through the Indigenous PhD fellowship program and International Research Support Initiative Program (IRSIP). I am thankful to Dr. Ghazala Kaukab, Chairperson, Department of Biochemistry, and member of my supervisory committee for her valuable scientific guidance and suggestions.

I would like to thank my co-supervisor, Dr Tariq Mahmood, Chairperson, Department of Plant Sciences, Quaid-i-Azam University, for his valuable suggestions and indebted help during the course of my work.

I am deeply grateful to the second member of my supervisory committee Prof. Dr. Tariq Masud for his practical and theoretical advices in the completion of this work.

I would like to thank Prof. Karen Skriver for giving me the opportunity to work in her group, for her continued guidance, support and interest in my varied
endeavors. I am highly indebted to Dr. Charlotte O’Shea for fruitful research discussion and always offering her help dealing with living in Denmark. I am thankful to all the lab members of Karen’s group, especially Maria, Anita, Lise, Sidsel and Nicolai, for making my stay wonderful in the Department. I am very grateful to my dearest friend Mahwish Khan and her family for their constant support and love during my stay in Denmark.

I wish to express my heartfelt thanks and obligations to my friends Sidra Younis, Safia Janjua, Shumaila Naz, Tehseen Rubbab, Fariha Khan, Bushra Javaid, Tasawar Sultana, Salma Batool, Maria Mushtaq, Amna Muhammad and Umara Afzal for their encouragement, support and company. I am grateful to all my colleagues and friends from plant Biotechnology Lab-1.

Last but not the least; I would like to express my deep gratitude to my parents for their constant love and support, being confident in me all through these years. I would like to mention my special gratitude to my mother without whom I would never come so far to pursue my Ph.D. I am so thankful to my brothers and sisters for all their love and giving me the strength to achieve my goals. I am cordially indebted to my husband, Muhammad Umar Younis, for his love, support and encouragement.

May Almighty Allah be their Gracious Guide and Source of Success and noble achievement in all their pursuits and in life (Ameen).

Farah Deeba
ABSTRACT

Germin and germin-like proteins (GLPs) belong to cell wall glycoproteins that have shown a significant resistance to detergent action, denaturation by heating, and degradation by proteases. GLPs expression is reportedly modulated during exposure to pathogens and abiotic stresses. Yet, little is identified about the regulatory mechanism of the GLP genes. The promoter of OsRGLP2 gene was isolated and cloned by Mahmood et al. (2007). This promoter showed strong expression of the GUS gene in transgenic tobacco during salinity, dehydration and wounding stresses. In this study, the regulatory cis-elements and their binding proteins for OsRGLP2 promoter are characterized. Various putative stress responsive cis-regulatory motifs and their specific binding proteins were identified by in silico analysis. DNA binding domains of OsWRKY71, OsDOF18 and OsMYB1 were cloned, overexpressed in E. coli and then purified by affinity chromatography using Glutathione Sepharose resin followed by cationic exchange chromatography. Electrophoretic mobility shift assays (EMSAs) have shown that recombinant OsWRKY71, OsDOF18 and OsMYB1 proteins were capable to interact with DIG labeled fragments of OsRGLP2 promoter containing W-box, AAAG and WAACCA motifs respectively. Binding was further confirmed by competitor EMSA and EMSA with mutant oligonucleotides. These regulatory elements were also active in binding with nuclear factors from rice nuclear proteins extract in vitro as confirmed by competitive EMSA. Expression analysis was performed to check the level of OsWRKY71, OsDOF18 and OsMYB1 against salt, cold, heat, wounding and drought stresses. Expression of OsWRKY71 was found to be induced in case of salt and cold stresses while OsDOF1 was expressed at relatively high level in response to salinity and drought. OsMYB1 expression was 23 fold higher in response to wounding which demonstrates the value of OsMYB1 up-regulation in
response to wounding stress in rice. In order to investigate the function of OsWRKY71, OsMYB1 and OsDOF18 against diverse abiotic stresses, recombinant plasmids were subjected to transformation in E. coli and their effect on E. coli growth was analyzed. The E. coli cells containing pGEX-OsWRKY71, pGEX-OsDOF18 and pGEX-OsMYB1 has shown different levels of tolerance against salt, drought, cold and heat stresses as compared to empty pGEX-4T1 vector. *In silico* characterization suggested the involvement of these proteins in protein-protein interaction. In conclusion, the positive response of OsWRKY71, OsDOF18 and OsMYB1 in abiotic stresses suggests their association with OsRGLP2 promoter and the importance of these proteins in providing protection to plants under abiotic stresses. Overexpression of OsWRKY71, OsDOF18 and OsMYB1 genes in crop plants may help in obtaining stress tolerant lines.
INTRODUCTION

Several biotic and abiotic factors, for example wounding, salinity, pathogen infection, drought and cold reduce plant growth rate as well as productivity. In response to these stresses, plants molecular systems co-evolved and consequently altered the molecular, biochemical and physiological responses to adopt stress environment (Bohnert et al., 1995). Many stress response genes and regulatory elements have been discovered by multiple research groups (Bray et al., 2000; Zhu, 2002). These molecules modify cell gene expression to alleviate cell damage caused by different stresses (Xiong and Yang, 2003; Jayasekaran et al., 2006). The Germin-like proteins (GLPs) that provide resistance to biotic and abiotic stresses by strengthening the cell wall belong to a distinct and a large group of plant stress-responsive genes. Importantly, GLPs protect the cell wall from proteases mediated degradation and also from heat- and detergent-responsive denaturation (Vallelian et al., 1998; Bernier and Berna, 2001). These proteins are named “GLPs” due to sequence and structural similarity with cereal germin proteins. Both, GLPs and Germins share profoundly preserved β-barrel and thus grouped into functionally divergent cupin superfamily (Dunwell et al., 2000).

GLPs are broadly distributed among different plants for instance, in A. thaliana twenty seven GLPs (El-Sharkawy et al., 2010) and in O. sativa about thirty GLPs have been discovered so far (Manosalva et al., 2009). The presence of a cleavable signal sequence in GLPs signify extracellular localization of these proteins (Carter and Thornburg, 1999). The GLPs possess diverse functions due to their degenerate active site. The GLPs and Germins execute four distinct enzyme functions; first true Germins mediate oxalate oxidase (OXO) activity (Berna and Bernier, 1997); which has
also been discovered in a GLP recently (Sakamoto et al., 2015). Second activity has been mediated by a number of GLPs called the superoxide dismutase (SOD) activity (Yamahara et al., 1999; Woo et al., 2000). Third, ADP glucose phosphodiesterase (AGPPase) activity and finally polyphenol oxidase (PPO) activity have also been predicted to be executed by GLPs (Rodriguez-Lopez et al., 2001, Cheng et al., 2014). Recently a GLP, Peruvianin-I, has been identified in the latex of Thevetia peruviana exhibiting cysteine peptidase activity (de Freitas et al., 2016).

The expression of GLPs is regulated by development and stress specific stimuli. The GLPs are expressed at specific developmental phases such as embryogenesis (Neutelings et al., 1998), floral stimulation (Staiger et al., 1999), and expansion of secondary xylem arrangement (Allona et al., 1998). The expression of GLPs and many Germins is also upregulated by abiotic stimulus, for instance, salt stress (Nakata et al., 2002; Jiang et al., 2007) and biotic stimulus such as pathogen infection (Dunwell et al., 2000; Park et al., 2004). Interestingly, in genetically modified plants GLP overexpression has been associated with enhanced resistance to salinity as well as pathogen stress (Lu et al., 2010; Knecht et al., 2010; Wang et al., 2013) and GLPs SOD activity was supposed to be responsible for providing this resistance (Christensen et al., 2004).

The gene expression is necessarily regulated in all living organisms. Similarly, in plants a number of the regulatory strategies have been identified which contribute in host defense response against biotic as well as abiotic stresses (Seki et al., 2001) and also modify the growth processes. The gene expression is principally regulated at transcription level and specifically controlled by cis-acting elements. Sequence specific DNA-binding proteins interact with promoter and regulate the transcription process through interaction with these cis-regulatory elements and also with enhancer
section of relative gene. The interaction of these proteins, i.e. transcription factors with correct *cis*-acting elements promote the binding of other proteins thus resulting in the formation of transcription complex. Ultimately, stress-response signaling and regulatory pathway genes are upregulated. In less developed organisms, a single transcription factor may be enough for transcriptional regulation, however, in highly developed organisms, appropriate complex of many transcription factors is necessary for transcription initiation and regulation (Lloyd *et al.*, 2001). In support, approximately 7% coding region of plant genome is encoded for transcription factors. It is necessary to discover transcription factors and corresponding binding locations on gene to identify the signaling pathways involved in regulation of a particular gene expression.

Advanced molecular techniques have assisted to identify and characterize specific protein quickly and precisely. In this regard, gel retardation assays or band shift assays which are commonly recognized as electrophoretic mobility shift assays (EMSAs) have been established to be a quick and accurate technique for the detection of distinct DNA-protein binding. This technique was first presented in 1981 by two independent research groups (Fried and Crothers, 1981; Garner and Revzin, 1981). The principal of this method depends on mobility retardation of the labeled oligonucleotide probe on non-denaturing polyacrylamide gel after binding with a sequence specific transcription factor.

GLPs and Germin are repetitively identified in plants under drought (Li *et al.*, 2010), wounding (Tabuchi *et al.*, 2003) and salinity (Nakata *et al.*, 2002) stresses. The characterization of promoter sites and relative transcription factors can be helpful to understand the function of GLPs under stress conditions. The promoter specific controlling *cis*-elements and their associated DNA binding proteins assist in providing
crucial terminals for cell signaling pathways that sensitize a cell to acclimatize to developmental or ecological changes. For instance, GER4c gene promoters were overexpressed in response to pathogen assault in barley (Himmelbach et al., 2010) and GLP13 was overexpressed in A. thaliana in response to oxidative stress (Tang et al., 2011).

A number of biotic as well as abiotic stress-response cis-regulatory elements have been recognized in various plants (Higo et al., 1999; Singh et al., 2002; Rani, 2007). The GLP promoters have been analyzed in silico by many research groups. For example, Sassaki et al., (2014) investigated EgGLP and identified cis-elements associated with auxin, light, Salicylic acid and Abscisic acid (ABREs) sensitivity. In addition, among seven GLP promoter sites from various plants, twelve regulatory cis-elements were commonly identified, such as ARR1, AAAG, W-box as well as GT and ACGT elements. The copy number of these cis-elements is plant specific (Mahmood et al., 2007). Recently another group identified many GER4c promoter specific cis-elements and analyzed seven W-boxes. Interestingly, they found that a W-box is necessary for stimulation by salicylic acid and functions in pathogen attack (Himmelbach et al., 2010).

Various transcription factors have been identified in different plants which regulate stress-response gene expression by interaction with corresponding regulatory elements. The WRKY proteins are among one of such proteins which regulate stress-specific genes expression during stress situations in plants (Zheng et al., 2006). Stress-response, flower formation and seed development is regulated by bZIP transcription factors which bind to ABREs cis-acting elements (Casaretto and Ho, 2003). Furthermore, abiotic stresses, e.g., drought, salt and abscisic acid induce a diverse MYB family specific gene expression regulators in many plants (Abe et al., 2003). In
response to drought, stimulation of dehydration responsive element binding (DREB) contributes in regulation of stress-response genes expression in plants through interaction with the promoter specific A/GCCGAC motifs (Sakuma et al., 2002). Moreover, the DNA-binding with one finger (DOF) proteins also regulate gene expression during plant development in addition to gene regulation in stress conditions (Chen et al., 1996; Kang et al., 2003).

In a study on GLPs, the characterization of OsRGLP2 identified five common promoter sequences namely CR1, CR2, CR3, CR4 and CR5. These sequences harbor various elements critical for dark and light response, senescence, salinity, plant growth regulators, pollen specific expression, pathogen infection and also comprise elements crucial for seed storage-proteins (Mahmood et al., 2007). Importantly, this promoter has also shown robust expression of a GUS gene in tobacco plants during various abiotic stresses for instance dehydration, salt, mechanical/ wounding stress (Mahmood et al., 2013). This finding suggests that the OsRGLP2 promoter is the site where various signaling pathways are converged to initiate a cascade of complex responses to stress-stimuli.

Further investigation into the importance of OsRGLP2 promoter under stress conditions through 5’deletion of OsRGLP2 promoter revealed that full length OsRGLP2 promoter as well as two fragments of sizes 1063 bp and 776 bp were sensitive to temperature, salt and wound stresses. However, additional deletion produced a fragment of 283 bp that had shown no response to treatment with various stress conditions. Therefore, a fragment of the OsRGLP2 promoter of size 565 bp may prove an efficient alternative of native promoter by the virtue of its smaller size and efficient activity against various abiotic stress stimuli (Hussain, 2015).
While the GLPs from different plants have already been characterized (Yin et al., 2009; Banerjee et al., 2010; El-Sharkawy et al., 2010), the regulatory means of their expression is still ambiguous. The characterization and further analysis of OsRGLP2 promoter may assist to identify specific transcription factors that could be a valuable point of entrance into the signaling cascade(s) between stress stimuli and the plant’s stress specific response. The characterization of GLPs transcription factor will not only provide us a comprehensive understanding of regulatory pathways, but also will give us chance to modify them in a beneficial way to meet continuously growing demand of increased food production.

The current study concentrated on the following aims to further explore the molecular pathways central for transcriptional regulation of the OsRGLP2 gene:

1. Identification of stress responsive cis-regulatory elements in the promoter region of OsRGLP2 gene and their specific binding proteins by in silico analysis

2. Cloning and overexpression of identified target proteins in E. coli expression system

3. Purification of target proteins from E. coli

4. Characterization of identified proteins by
   - Binding studies with OsRGLP2 promoter fragments using Electrophoretic Mobility Shift Assays (EMSAs)
   - Expression analysis under different abiotic stresses by quantitative Real time PCR (qRT-PCR)
   - Abiotic stress analysis in E. coli
   - In silico characterization for protein-protein interaction studies
REVIEW OF LITERATURE

2.1 GERMIN AND GERMIN LIKE PROTEINS (GLPs): THE DIVERSE MEMBERS OF CUPIN SUPERFAMILY

Germin and GLPs are classified as component of the functionally diverse proteins family, called cupin superfamily, due to the occurrence of a common, extremely preserved β-barrel (Dunwell et al., 2000). This superfamily includes proteins from Archaea, Eubacteria, and Eukaryota (Khuri et al., 2001) having both enzymatic and nonenzymatic functions (Membre et al., 2000). Other members of the cupin superfamily contain auxin-binding proteins, cyclases and dioxygenases that are proposed to play role in abiotic stresses such as heat and salt stress and aluminum exposure.

Germin is an apoplastic, glycosylated protein marker, possesses oxalate oxidase (OXO) activity and was first identified in germinating wheat embryos (Thompson and Lane, 1980). Mature germin protein is a homohexamer comprising of six β-jellyroll monomers which accounts for its remarkable resistance to the actions of proteases, heat and detergent (Lane et al., 1993; Membre et al., 2000). GLPs are heterogenous proteins closely related to germins, sharing 25%-100% sequence homology among themselves and 30%-70% with germins. GLPs are disseminated among monocots, dicots, angiosperms, gymnosperms and bryophytes (Bernier and Berna, 2001).

2.2 BIOCHEMICAL NATURE OF GLPs

Germin and GLPs are found to participate in protein-protein interaction. Both germins and GLPs possess various types of enzymatic activities.
2.2.1 Oxalate Oxidase (OXO) Activity

Oxalate oxidase activity is linked with true germin proteins (Berna and Bernier, 1997). In contrast to germins, oxalate oxidase activity had not been associated with any of GLPs identified but recently Sakamoto et al. (2015) reported a GLP from woody azalea possessing oxalate oxidase activity.

2.2.2 Superoxide Dismutase (SOD) Activity

Most of GLPs possess SOD activity (Woo et al., 2000). Superoxide dismutase (SOD) is responsible for producing hydrogen peroxide ($\text{H}_2\text{O}_2$) from superoxide radical ($\text{O}_2^-$) and consequently serves as a critical enzyme in providing protection to plants against reactive oxygen intermediates.

2.2.3 ADP Glucose Pyrophosphatase or Phosphodiesterase (AGPPase) Activity

GLPs also exhibit phosphodiesterase or ADP glucose pyrophosphatase (AGPPase) activity (Rodriguez-Lopez et al., 2001). GLPs isolated from wheat apoplast caused inhibition of serine protease and its expression increased upon attack by microbes (Segarra et al., 2003; Cordo et al., 2007).

Two isoforms of AGPPases have been identified from barley leaves (Rodriguez-Lopez et al., 2001). By sequence analysis, it was found that these two isoforms are different oligomers of $H. \text{vulgare}$ GLP1. Additionally, a maize GLP1 is also reported to possess AGPPase activity having sequence similarity with barley GLP (Fan et al., 2005).

2.2.4 Polyphenol Oxidase Activity

Phenolic compounds accumulate in response to infection and are considered to be involve in cell wall reinforcement. Recently, a protein with high PPO activity was
identified as GLP in the peel fraction of citrus (Cheng et al., 2014). GLP catalyzes the oxidation of phenolic compounds to Quinones and results in browning.

2.3 REGULATION OF GLPs EXPRESSION DURING VARIOUS STRESSES

GLPs are involved in plant defense against pathogen attack because several GLPs possess SOD activity that results in production of hydrogen peroxide, which has been supposed to be a signaling particle for a variety of defense responses, for instance, cell death, and play a role in cell wall reinforcement by acting as a cofactor (Yamahara et al., 1999; Laloi et al., 2004).

2.3.1 Role of GLPs in Biotic Stresses

Germin and GLPs are well characterized for their role in providing protection against fungal attack. The expression of many GLP genes is activated by fungal and microbial infections (Schweizer et al., 1999; Godfrey et al., 2007; Davidson et al., 2009). After infection of rice with M. oryzae, GLP gene expression was induced (Davidson et al., 2009; Manosalva et al., 2009). GLP gene expression was observed to be induced after attack by the powdery mildew pathogen in barley and wheat (Wei et al., 1998; Zimmermann et al., 2006) GLP expression was enhanced by Fusarium graminearum attack in barley spikelets (Federico et al., 2006). GLP expression was enhanced in grapevine leaves and fruits after Erysiphe necator infection (Godfrey et al., 2007).

2.3.2 Role of GLPs in Abiotic Stresses

Several abiotic stresses have been shown to regulate the expression of GLPs. A GLP gene of moss Barbula unguiculata, BuGLP, has been found to be up-regulated in
cells in logarithmic phase in response to salinity (Nakata et al., 2002). In Atriplex lentiformis, salt stress and ABA treatment inhibited expression of AlGLP. Wounding and methyl jasmonate treatment caused increase accumulation of AlGLP mRNA in leaves (Tabuchi et al., 2003). Li et al. (2010) isolated a GLP from Tamarix hispida, ThGLP, and observed that its expression is up-regulated by drought, salinity, cold, Cadmium chloride (CdCl₂) and upon abscisic acid treatments.

2.4 GLPs PROMOTER ANALYSIS

A promoter is a fragment of DNA frequently occurring upstream from a gene coding region and acts as to regulate the expression of a gene. The promoter section consists of definite DNA sequences which contains regulatory elements that work in the conscription of proteins. These proteins make possible transcription of a gene. These regulatory cis-acting elements are also called transcription factor binding sites (TFBS) and these elements decide the spatiotemporal expression of a gene. The regulatory elements or TFBS are precise short stretches of DNA typically 5-25 bp (Rani, 2007) and transcription factors bind to these regulatory elements to turn on or turn off the gene transcription in response to internal or external environmental factors.

Promoters can be divided into three distinct types: constitutive promoters, tissue specific promoters or development stage specific promoter and inducible promoters. The interest in inducible promoter systems is increasing because it allows plants to self-regulate the gene expression under harsh internal or external environmental conditions.

Mathieu et al. (2003) cloned and checked the expression pattern of a 5’-upstream promoter region of 1520 bp Pinus caribaea germin1 (PcGLP1). The sequence of this promoter contains several cis elements for hormonal response and for
expression in embryo or during the process of germination. This promoter was cloned in fusion with the GUS (β-glucuronidase) reporter gene and was transformed into tobacco Bright Yellow 2 cells under different hormonal conditions. They found that 5’ promoter deletions affected the reporter gene action under optimal growth conditions.

Deletion analysis of 1376 bp ZmGLP1 promoter in A. thaliana showed that different deletion promoters have different expression intensity in transgenic plants (Fan et al., 2005). The strongest GUS expression was obtained in 739 bp promoter and mostly expression was in filaments and most of the green tissues.

Himmelbach et al. (2010) have shown that an intense cluster of tandemly duplicated genes is present in barley which have experienced some cycles of duplication. Various WRKY transcription factors binding sites are located in GER4 promoter region that were traded between subfamily individuals by gene conversion. Mutational analysis was carried out to disclose their embellishing effects and functional versatility on pathogen-activated promoter activity.

Sassaki et al. (2015) investigated the EgGLP expression pattern using a 1300 bp EgGLP promoter fused to a GUS reporter gene. GUS expression was observed in transient as well as in stably transformed tobacco plants. It was observed that activity was modulated by light/dark transitions.

2.5 CIS-ACTING REGULATORY ELEMENTS FOR ABIOTIC AND BIOTIC STRESSES

Different types of cis-acting regulatory elements were discovered in plants that play a role in response to specific or multiple stresses such as ABA responsive elements, W-boxes, AAAG motif, ABRE motif and ARR1.
2.5.1 Abscisic Acid Responsive Element (ABRE) Motifs

The ABRE sequence was first discovered in *T. aestivum* EM1A gene promoters PyACGTGGC (Guiltinan *et al.*, 1990). The core sequence of ABRE is ACGT, similar to the G-box which responds to various stresses such as UV light and osmotic stress. Research has revealed that single ABRE motif was not sufficient to excite the promoter activity in ABA treatment (Skriver *et al.*, 1991). After the identification of coupling elements (CEs), high level ABA induction was observed when ABRE and CE appeared at the same time (Shen and Ho, 1995). ABRE and CE elements are found prevailing in 1000 bp promoter regions of stress and ABA responsive genes.

2.5.2 Drought Responsive Elements (DREs) Motifs

A class of cis-acting regulatory elements receptive to salt stress, dehydration and cold stress is Drought responsive elements (DREs). This motif was first acquired from the promoter of the Arabidopsis *RD29A* gene (Yamaguchi-Shinozaki and Shinozaki, 2006). It was found that DRE and ABRE both are situated in the promoter section of the *RD29A* gene under osmotic stresses. DRE serves as a rapid-responsive cis-acting elements in first 20 minutes and ABRE exhibits a slow induction of gene expression in the ongoing hours (Yamaguchi-Shinozaki and Shinozaki, 1994).

2.5.3 Ethylene-Responsive Elements (EREs) Motifs

Ethylene regulates the processes of seed germination, senescence and stress factors such as flooding, wounding and pathogen attack (Bleecker and Kende, 2000). A functional ERE was identified from tomato E4 and E8 genes promoters (Montgomery *et al.*, 1993; Deikman, 1997). It has been found that as a minimum two cooperative cis-acting sequences are required for ethylene receptive regulation of E4 (Xu *et al.*, 1996). A nuclear protein that binds to the 5’-flanking regions of E4 and E8 genes was identified by gel shift assays and DNAse I footprinting.
2.5.4 Anaerobic-Responsive Elements (AREs) Motifs

In response to low oxygen, expression of a specific set of genes is induced. Deletion mapping identified a sequence homologous to GT motif T/CGGT/T and the GC motif GCCG/CC. The functional properties of the AREs in the Adh1 gene promoter have been analyzed (Dolferus et al., 1994).

2.5.5 W-Boxes

There is rising proof that main class of cis-regulatory element present in promoters of various plant genes that plays role against pathogen contamination and abiotic hassles is W-box. The significance of the W-boxes was demonstrated lately by analysis of the Arabidopsis transcriptome in SAR (Maleck et al., 2000).

2.6 TRANSCRIPTION FACTORS ARE INVOLVED IN BIOTIC AND ABIOTIC STRESSES

Transcription factors are critical regulatory proteins which bind to definite promoter sequences upon actuation or de-enactment of upstream signaling network and turn on or turn off the transcription of the target genes. The transcription factors are made up of minimum two distinct domains: a DNA binding domain, which makes interaction with cis-regulatory elements located in the target gene promoters and an activation domain which activate the transcription of target genes or repression domain which turn off the transcription of particular genes. Plant genomes devote 7% of their protein encoding genes to transcription factors (Udvardi et al., 2007) and several of these genes provide instant response to abiotic stresses. These transcription factors make interaction to their respective cis-regulatory elements located in the promoters of specific genes and in turn up- or down-regulate the expression of several downstream target genes which results in providing abiotic stress tolerance.
Recent researches recognized numerous transcription factors vital in controlling plant behavior to diverse stresses (Bohnert et al., 2001; Seki et al., 2001; Zhu et al., 2005; Agarwal et al., 2006). Transcription factors have been clustered into different families because of the different structure of DNA binding domains. Out of numerous transcription factors families, Myeloblastosis viral oncogene homolog (MYB), Myelocytomatosis viral oncogene homolog (MYC), basic leucine zipper (bZIP) ethylene response factor (ERF) and WRKY are involved in responding to a variety of stresses (Schwechheimer et al., 1998; Singh et al., 2002) and DNA binding with one finger transcriptional factors (DOF) are implicated in the expression of developmental genes for instance, genes involved in photosynthesis, seed storage and genes activated by plant hormones and various stress indications (Baumann et al., 1999; De Paolis et al., 1996; Kisu et al., 1998).

2.6.1 WRKY transcription factors

WRKY proteins constitute one of the major groups of plant specific transcription factors that control several processes in plants. First report about WRKY transcription factors came twenty years ago and since then WRKY transcription factors have achieved a considerable progress. There are several reports regarding the functionality of WRKY proteins in providing defense against stresses, seed germination and regulation of diverse developmental stages in plants (Rushton et al., 2010).

2.6.1.1 WRKY domain and classification

The WRKY transcription factors are named because of the DNA binding domain, the major feature of these transcription factors. This domain is known as WRKY domain due to the occurrence of the invariant WRKY heptapeptide (WRKYGQK) also designated as the “signature sequence” present at the N-terminus
of DNA binding domain (Maleck et al., 2000). The WRKY DNA binding domain is composed of 60 amino acid residues and consists of two parts. The WRKY signature sequence is present at the N terminus side and a typical zinc finger domain is located at the C-terminus. The zinc-finger domain is either C–X_{4.5}–C–X_{22–23}–H–X_1–H or C–X_7–C–X_{23}–H–X_1–C depending upon the group of WRKY genes to which the protein belongs.

Initially, WRKY family was clustered into three subfamilies because of the presence of more than one WRKY domain in the protein and the structure of the associated Zinc finger motif. Group I contains proteins having two WRKY domains, while the majority of proteins belong to Group II and III containing one WRKY domain. However, the similar type of Zinc finger domain (C–X_{4.5}–C–X_{22–23}–H–X_1–H) is present in members of group I and II. C2–HC type zinc finger motif (C–X_7–C–X_{23}–H–X_1–C) is found in group III WRKY domains. Later, detailed phylogenetic analyses revealed that WRKY family can be more precisely separated into I, IIa+b, IIc, IId+e and III in higher plants (Rushton et al., 2010; Zhang and Wang, 2005).

### 2.6.1.2 WRKY transcription factors in stresses

Numerous reports have confirmed that WRKY transcription factors participate in a complex regulatory network against many stresses in plants such as wounding, drought, high temperature, low temperature or salinity. ABA is known as the stress hormone as it conciliates plant responses to abiotic stresses. ABA biosynthesis is raised because of drought, cold and salt stresses. In Boea hygrometrica, BhWRKY1 control the expression of BhGolS1 and directs to improve drought tolerance (Wang et al., 2009). Besides, AtWRKY33 transcript level is increased in response to salt stress in an ABA dependent manner. However, AtWRKY25 transcript accumulation is ABA
independent (Jiang and Deyholo, 2009). Overexpression of AtWRKY25 and AtWRKY33 improved salt tolerance, but resulted in more responsiveness to ABA and oxidative treatments. In wheat, 8 of 15 WRKY genes were also found sensitive to temperature stress, salt or PEG treatment (Wu et al., 2009).

Numerous WRKY proteins were observed to be concerned in plant in drought and salt stresses. In rice, overexpressing OsWRKY11 downstream of heat shock inducible promoter HSP101 caused increase tolerance to drought and heat stress (Wu et al., 2009). Similarly, overexpression of OsWRKY45 also improved tolerance to salinity and drought (Qiu and Yu, 2009). Zhou et al. (2008) discovered that overexpression of GmWRKY21 appeared to have more tolerance in transgenic plants against cold stress as compared to untransformed control plants. While plants overexpressing GmWRKY54 and GmWRKY13 revealed to have more tolerance against salinity and drought stresses (Zhou et al., 2008).

2.6.2 MYB Transcription Factors in Plants

v-myb was first MYB gene recognized from Avian myeloblastosis virus from which name MYB was derived (Klempnauer et al., 1982). The occurrence of one to four MYB repeats (R) form the major feature of all MYB proteins and these MYB repeats can work collegially or independently in DNA binding and also in interaction with other proteins. The length of a MYB repeat is about 52 amino acid residues and every repeat includes three conserved tryptophan residues which are regularly spaced and are involved in forming a hydrophobic core (Kane-Ishii et al., 1990). Each MYB repeat endorses a helix-turn-helix conformation, (Ogata et al., 1996). The third α-helix in each MYB repeat is involved in binding with DNA major groove (Ogata et al., 1996; Jia et al., 2004).
2.6.2.1 DNA-binding specificity of MYB factors

Recognition helices of MYB proteins dimerize and bind with DNA (Sakura et al., 1989). Plant MYB DNA target sequence was first resolved for Maize P protein, which is an R2R3 type MYB protein implicated in flavonoid biosynthesis (Grotewold et al., 1994). Through EMSAs and binding site selection assays have shown that P protein binds with ACC(A/T)ACC(A/C/T). Different MYB proteins identify different cis-acting regulatory elements.

2.6.2.2 Role of MYB in abiotic stresses

The R2R3 type MYB family is greatly extended in plants (Dubos et al., 2010), and numerous MYB genes are involved in abiotic stress responses. MYB44/MYB91 overexpression has caused the repression of JA-sensitive gene activation. MYB96 regulated drought stress responses in A. thaliana by assimilating signals from auxin and ABA (Seo et al., 2009). MYB2 expression was brought about by ABA and dehydration. When MYB2 was overexpressed in Arabidopsis, transgenic Arabidopsis plants showed hypersensitivity to ABA and improved response to osmotic stress in comparison to wild type Arabidopsis (Abe et al., 2003). MYB15 recognized, bound to MYB- binding sequence in DREB1A, DREB1B and DREB1C promoters and resulted in down-regulation of these genes. MYB15 was also linked to other abiotic stress responses, as MYB15 overexpression resulted in increase tolerance in Arabidopsis against drought, freezing and salt stress in (Agarwal et al., 2006; Ding et al., 2009).

2.6.3 DNA With One Finger (DOF) Transcription Factors

The DNA-binding with one finger (proteins) transcription factors are a group of plant specific transcription factors. DOF proteins have been analyzed from rice, Arabidopsis, soybean and poplar (Yanagisawa, 2002; Lijavetzky et al., 2003; Yang et al., 2006). The DOF proteins contain Cys/Cys type zinc finger DNA binding domain
which is approximately 52 amino acid residues long (Yanagisawa, 1995). The DOF domain has been identified in various DOF proteins, including maize DOF1, DOF2 and PBF (Yanagisawa, 1995; Vicente-Carbajosa et al., 1997; Yanagisawa and Sheen, 1998).

The zinc finger domain is able to bind to AAAG motif in the target gene promoter region (Yanagisawa, 1995) and is also involved in protein-protein interaction including interaction with DOF itself (Yanagisawa, 1997; Diaz et al., 2002). DOF proteins are classified into different groups based upon DNA-binding zinc finger domain. Difference in DNA binding zinc finger domain determines the binding affinity of DOF with AAAG motif.

2.6.3.1 Role of DOF in stresses

DOF proteins have been described to participate in controlling diverse processes as seed germination (Yamamoto et al. 2006), tissue specific gene expression (Kim et al., 2010), flower abscission (Wei et al., 2010) and plant hormone signalling (Moreno-Risueno et al., 2007). The expression level of DOF proteins is also regulated by various types of abiotic and biotic stresses. The expression of many DOF proteins is induced by various phytohormones and by pathogens attack (Nakano et al., 2006; Skirycz et al., 2007). Moreover, the promoters of pathogen-related elicitors genes frequently include putative DOF binding sites AAAG motif (Kang et al., 2003; Haque et al., 2009; Gomez-Ros et al., 2012). OBP2 is implied in the biosynthesis of indole glucosinolate in A. thaliana. OBP2 is tissue-specific and its expression is inferred in reaction to herbivore and by methyl jasmonate hormone.
2.6.4 NAC (NAM/ATAF1/CUC2) Transcription Factors

NAC (NAM/ATAF1/CUC2) proteins are transcription factors, which have so far been found only in land plants. They play important roles in development, abiotic and biotic stress responses, and biosynthesis processes (Nakashima et al., 2012; Puranik et al., 2012). Their DNA binding domain, the NAC domain, interacts with the CGT (AG) core sequence of target genes (Jensen et al., 2010), and its structure and DNA-binding mode suggest an evolutionary relationship with WRKY and GCM transcription factors (Babu et al., 2006). The specific molecular interactions of NAC proteins with specific genes and other proteins are essential for regulation of networks with NAC proteins and genes as central components. The network of Arabidopsis ANAC019, ANAC055, and ANAC072 regulates both abiotic stress responses and pathogen sensitivity, and in addition the corresponding emerging network for rice reveals epigenetic regulation of growth (Tran et al., 2004; Jensen et al., 2010; Nuruzzaman et al., 2010). Knowledge of NAC protein structure-function relationships and networks is essential for future application of NAC proteins in agriculture.

2.7 STRATEGIES TO UNRAVEL THE FUNCTION OF PLANT PROMOTERS

The better understanding of cis-acting regulatory elements and their corresponding binding proteins can help us to interpret the regulatory activities of promoters. Using known cis-acting regulatory elements, responding transcription factors can be identified (Lopato et al., 2006). Numerous methods have been used to identify and characterize the protein binding to DNA in vitro such as Electrophoretic mobility shift assays (EMSAs), DNase I footprinting, and filter binding assays.

2.7.1 Filter Binding Assays

Filter binding is easy to carry out and the exploitations are fast to permit kinetic
studies in addition to equilibrium dimensions (Riggs et al., 1970; Woodbury et al., 1983). Filter binding is a nonequilibrium method similar to the EMSA. This method is not affected by the concentration of sample salt and it can have room for large nucleic acids molecules (Whitson et al., 1986). The existence of more than one protein makes filter-binding analysis difficult, because retention of labeled nucleic acid is analyzed and not bound proteins or the measure of binding activity. For this reason, this assay is not suitable for discriminating one protein-containing complexes from those containing multiple proteins.

2.7.2 DNase I Footprinting

DNase I Footprinting assays make use of the observation that a nucleic acid bound protein will hinder with the enzymatic or chemical alteration of that nucleic acid sequence (Brenowitz et al., 1987). Many chemicals are used for this alteration. In this method, one nucleic acid strand label with a radioisotope is positioned at one end. After modification and cleaving the nucleic acid at a modification site if required, the reaction is run on a denaturing polyacrylamide gel. Protected sequences are then identified by comparing reaction patterns of products. Additionally, the manifestation of sites that are hypersensitive to alteration can offer proof of modification of the target nucleic acid sequence (Tullius, 1989). Because sequences of hundreds of amino acids can be set on a classic acrylamide gel, the method is suitable for the investigation of the multiple proteins bound to a single nucleic acid molecule. Some proteins interact with nonspecific DNA may not result in specific footprints. In these cases, signals obtained from EMSA and filter binding are more straightforward to construe than signals obtained from DNase I footprinting.

2.7.3 Electrophoretic Mobility Shift Assays (EMSAs)

Electrophoretic mobility shift assays (EMSAs) are widely used to scrutinize the
capability of a protein to interact with a DNA fragment and to measure the quantity of DNA bound to protein (Hudson et al., 1990). EMSA can detect even when protein concentration is lower in the extract. This is due to the binding of protein to only few probe molecules which results in complex formation and slow migration on gel as compared to free probe. EMSA can be used for a range of proteins (Talanian et al., 1990; Mita et al., 1995). It works well with both crude nuclear and cell extract and purified proteins (Varshavsky, 1987). This capacity accounts in huge for the ongoing fame of this method.

To comprehend transcriptional mechanisms that direct to differential gene expression, genes have to be experimentally synchronized with the consequent transcription factors that control their expression. 2025 putative transcription factors have been identified in Oryza sativa L. ssp. Indica and 2384 in ssp. japonica by manual and computer prediction (Gao et al., 2006). DRTF database contains information about the sequence, functional domains chromosomal location, and expression sequence tags (EST) and microarray expression information about transcription factors. Combinatorial regulation of GLPs promoters and their specific transcription factors is an essential area to understand the process of plant gene regulation but not much attention has been paid towards it. To increase the expression of genes engaged in stress conditions in a greatly synchronized and reversible approach, identification and characterization of GLPs promoter elements and corresponding transcription factors is required. It will help in developing stress tolerant transgenic lines for various important crops.
Chapter 3

MATERIALS AND METHODS

3.1 SEQUENCE ANALYSIS OF OsRGLP2 PROMOTER

OsRGLP2 promoter sequence was gained from NCBI database (http://www.ncbi.nlm.nih.gov/). OsRGLP2 promoter region of 1104 bp (Accession no. DQ414400.1) was subjected to different bioinformatics tools for the existence of putative cis-regulatory elements vital for recruiting stress responsive transcription factors to the promoter. PLACE online database (http://www.dna.affrc.go.jp/PLACE/signalscan.html) was used first to scan promoter and then ConSite (http://consite.genereg.net/), JASPAR (http://jaspar.binf.ku.dk/) and Plant Care (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) were used to scan promoter to further appraise the stress receptive cis-regulatory elements.

3.2 SEQUENCE RETRIEVAL AND IN SILICO ANALYSIS OF PROTEINS THAT PUTATIVELY BIND TO OsRGLP2 PROMOTER

Sequences (nucleotide and protein) of chosen stress related transcription factors were acquired from NCBI database (http://www.ncbi.nlm.nih.gov/). Then, BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was explored with retrieved protein sequences of MYB1 from Arabidopsis thaliana (AtMYB1) and DOF1 from Zea mays (ZmDOF1). BLAST search was narrowed to merely take account of hits from Oryza sativa genome against the non-redundant database. Nucleotide and protein sequence of OsWRKY71 obtained from NCBI was directly used for further in silico analysis. Homologous protein sequences from different plants were retrieved from BLAST outcome and multiple sequence alignment was executed using online TCOFFEE software (http://tcoffee.crg.cat/) to recognize the conserved sequences of DNA binding domains. Multiple sequence alignment
representation was produced using the ESPript online tool (http://escript.ibcp.fr/ESPript/ESPript/). PSIPRED online tool (http://bioinf.cs.ucl.ac.uk/psipred/) was employed to anticipate the secondary structure of DNA binding domains. Conserved DNA binding domains of OsWRKY71, OsDOF18 and OsMYB1 were calculated from SMART online tool (http://smart.embl-heidelberg.de/). Expasy’s ProtParam Proteomic server (http://web.expasy.org/protparam/) was used for computing basic physiochemical properties of OsWRKY71, OsDOF18 and OsMYB1 based on theoretical isoelectric point (pI), instability index, molecular weight, aliphatic index and grand average of hydropathicity (GRAVY).

3.3 CLONING OF PROTEINS THAT BIND TO OsRGLP2 PROMOTER

3.3.1 Primer Designing

To clone OsWRKY71, OsDOF18 and OsMYB1 DNA binding domains with additional 15-20 amino acid adjoining sequences N and C terminal, forward and reverse primers were designed with BamHI and XhoI restriction sites respectively. An effort was made to clone the full open reading frame (ORF) as well as only DNA binding domain of OsDOF18 gene and for this purpose, forward and reverse primers were designed with BamHI and XhoI restriction sites respectively. Primer sequences and their properties are mentioned in table 1. Their properties were checked using different bioinformatics tools i.e. Integrated DNA technology oligoanalyzer (www.idtdna.com) or Primer BLAST (www.ncbi.nlm.nih.gov/primerBLAST).

3.3.2 Plant Material

Seeds of Oryza sativa cv. KS282 were gained from Rice Program, Crop Science Institute, National Agricultural Research Centre (NARC) Islamabad.
rice seeds were placed in growth room for 10 days at 25°C after culturing on half strength MS basal medium (Appendix 1) (Murashige and Skooge, 1962).

3.3.3 RNA Isolation and Synthesis of cDNA

Total RNA was extracted from 10 days old rice seedlings using the RNeasy mini kit from QIAGEN according to the instructions provided with certain modifications. Briefly, 25-75 mg of both control and treated plant material was crushed in precool mortar and pestle with the help of liquid nitrogen. Then liquid nitrogen was allowed to evaporate and ground material was transferred to an eppendorf tube containing RLT buffer (450 µL) followed by vortexing. The lysate was then poured onto purple spin column, kept in collection tube and centrifugation was done at 10,000 X g for 2 minutes. After centrifugation, supernatant was added carefully to new eppendorf tube and 96% ethanol was added to it. Then the lysate was transferred to pink spin columns, placed in collection tubes and spin down at 10,000 X g for 1 minute. Run out was removed and of RW1 buffer (350 µL) was poured in the column and spin down again at the same speed for 1 minute. Then DNAse I solution was poured on column and kept at room temperature for 20 minutes to facilitate DNA digestion. After that, RW1 buffer (350 µL) was transferred to the column and spinned down at the same speed for 1 minute. Run out was removed and 500 µL of RPE buffer was applied on column to clean the membrane followed by centrifugation at same speed for 2 minutes. After discarding flow through, spin column was kept in a new collection tube and spin down at the same speed for 1 minute. Again new collection tube was placed and RNAs free water (50 µL) was added carefully on top of the column and spinned down at same speed for 1 minute. Isolated RNA was stored at -70°C.
Table: 1: Primer sequences used for cloning of *OsWRKY71*, *OsDOF18* and *OsMYB1* DNA binding domains and full length *OsDOF18*

<table>
<thead>
<tr>
<th>Primers Name</th>
<th>Primers Sequences</th>
<th>Product size (bp)</th>
<th>Accession Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>OsWRKY71</em>(167-271)-For</td>
<td>5'-CGCGGGATCCATGCAGATCGCGAGGAG-3' (BamHI)</td>
<td>315</td>
<td>AY676927.1</td>
</tr>
<tr>
<td><em>OsWRKY71</em>(167-271)-Rev</td>
<td>5'-CCGCTCGAGTCAGATCGGGCGCTTTGCGAGGAG-3' (XhoI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>OsDOF18</em>(1-277)-For</td>
<td>5'-CGCGGGATCCATGCAGAGGAGCAGAGCGG-3' (BamHI)</td>
<td>831</td>
<td>NM_001068645.2</td>
</tr>
<tr>
<td><em>OsDOF18</em>(1-277)-Rev</td>
<td>5'-CCGCTCGAGTCATGGGAGGTTGAGGAGAAGAC-3' (XhoI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>OsDOF18</em>(1-148)-For</td>
<td>5'-CGCGGGATCCATGCAGAGGAGCAGG-3' (BamHI)</td>
<td>444</td>
<td>NM_001068645.2</td>
</tr>
<tr>
<td><em>OsDOF18</em>(1-148)-Rev</td>
<td>5'-CCGCTCGAGTCAGCTCCGGAGTCGTCAGG-3' (XhoI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>OsMYB1</em>(88-222)-For</td>
<td>5'-CGCGGGATCCATGCAGAGGAGCAGTAGTCTAC-3' (BamHI)</td>
<td>405</td>
<td>NM_001051349.1</td>
</tr>
<tr>
<td><em>OsMYB1</em>(88-222)-Rev</td>
<td>5'-CCGCTCGAGTCAGACCTTCTTGAGCAGG-3' (XhoI)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1.2% agarose gel (prepared with RNase free 1X TAE buffer) was run with isolated RNA samples to examine the quality and integrity of RNA. Loading dye was mixed with RNA samples, heated at 70°C for 3 minutes and kept on ice immediately. Electrophoresis was performed at 66V for 30 minutes. After gel running, it was exposed to UV for band visualization. Total RNA concentration was measured using nanodrop by taking absorbance at 260 nm.

Then cDNA was synthesized using using RevertAid Premium Reverse Transcriptase (Thermoscientific) by means of Oligo (dT) 18 Primer which were synthesized commercially.

3.3.4 PCR Amplification of OsWRKY71 and OsMYB1 DNA Binding Domains

DNA binding domains of OsWRKY71 and OsMYB1 were amplified using specific primers mentioned in table 1. PCR reaction mixture contained 4 µL of HF Phusion buffer (5X), 0.4 µL dNTPs (10 mM), 1 µL of both primers (10 µM each), 0.2 µL of Phusion DNA polymerase, 0.6 µL DMSO, 2 µL of cDNA and water up to 20 µL. To avoid 3’→ 5’ exonuclease activity of Phusion DNA polymerase, which causes primer degradation, enzyme was included in the reaction mixture at the end. Thermal profile included 1-initial denaturation by heating the PCR reaction (95°C for 1 minute), followed by 35 cycles of 2- denaturation (95°C, 20 seconds), 3-annealing (65°C for MYB1 and 66°C for WRKY71, 20 seconds) and 4-polymerization (72°C, 20 seconds) followed with a final polymerization step (72°C) of 10 minutes and maintenance of the samples at 4°C. After the addition of 3 µL of tracking dye, the PCR products were examined by loading the reaction mixture on agarose gel (1.5%) made in TAE buffer (1X) (Sambrook and Russell,
followed by electrophoresis at 90 V. For correct size estimation, gene ruler 1 Kb (Thermoscientific) was run on gel along with samples. Ethidium bromide stained gel was visualized by UV transilluminator. After cutting the specific sized fragment, the PCR product was purified according to manufacturer’s instructions using GeneJET Gel Extraction Kit (Thermoscientific).

### 3.3.5 PCR Amplification of OsDOF18

A cDNA clone J065152E11, which contains DOF18 coding sequence, was purchased from NIAS databank Japan. To use this clone for DOF18 cloning, lyophilized plasmid was resuspended first in Millipore water (10 µL) and then 1 µL was subjected to transformation in DH5α competent cells. Next day, starter culture was prepared from transformd colonies and plasmid was isolated using miniprep plasmid isolation kit (Thermoscientific). Isolated plasmid was subjected to restriction digestion for 1 hour at 37°C and confirmed by running on agarose gel. Restricted coding sequence was gel purified and then used as template in PCR reactions of DOF cloning (full length as well as DNA binding domain region).

PCR reaction mixture contained 20 µL of HF Phusion buffer (5X), 2 µL dNTPs (10 mM), 2 µL of both primers (10 µM each), 1 µL of Phusion DNA polymerase, 2 µL DMSO, 20 ng restricted fragment and water was added to make total reaction volume of 100 µL. Thermal profile included 1-initial denaturation by heating the PCR reaction (95°C for 5 minutes), then 25 cycles of 2-denaturation (95°C, 1 minute), 3-annealing (66°C for full length and 63°C for DNA binding domain, 1 minute) and 4-extension (72°C, 1 minute) followed with a final polymerization step (72°C) of 7 minutes and maintenance of the samples at 4°C.
After the addition of 3 µL of tracking dye, the PCR products were examined by loading the reaction mixture on agarose gel (1.5%) made in TAE buffer (1X) (Sambrook and Russell, 2001) followed by electrophoresis at 90 V. For correct size estimation, gene ruler 1 Kb (Thermoscientific) was run on gel along with samples. Ethidium bromide stained gel was visualized by UV transilluminator. After cutting the specific sized fragment, the PCR product was purified according to manufacturer’s instructions using GeneJET Gel Extraction Kit (Thermoscientific).

### 3.3.6 Cloning of *Os*WRKY71, *Os*DOF18 and *Os*MYB1 DNA Binding Domain Regions

pGEX4T-1 vector which contains Glutathione S-transferase (GST) gene as a tag was used to clone WRKY71 (167-271), DOF (1-277), DOF (1-148) and MYB1 (88-222). Fast Digest BamH1 and Xho1 restriction enzymes were used to digest pGEX4T-1 and purified PCR amplicon. Restriction digest Reaction consisted of fast digest buffer (4 µL), Fast Digest BamH1 and Xho1 (1 µL of each), 1000 ng pGEX4T-1 or 500 ng purified PCR product and water up to total sample volume of 40 µL. Samples were digested at 37°C for required time and then reaction was blocked by mixing 5 µL of loading dye.

To decrease the number of empty vectors and avoid circularization of plasmid during the ligation procedure, thermosensitive Alkaline Phosphatase was used to dephosphorylate the restricted pGEX-4T-1. For this purpose, volume of the restriction digest reaction was adjusted to 50 µL by the addition of Thermo Scientific FastDigest Buffer (5 µL), alkaline phosphatase enzyme (2.5 µL) and water (2.5 µL). The reaction was kept at 37°C for 1 hour. Purification of restricted
PCR product and dephosphorylated pGEX4T-1 was performed to remove the small sized fragments using PCR purification kit from Thermoscientific. To estimate the concentration of purified restricted PCR product and vector for ligation process.

In a ligation reaction, 1:3 ratio of vector to insert was used which was calculated by online available Ligation calculator (http://www.insilico.uni-duesseldorf.de/Lig_Input.html). For ligation, total reaction of 20 µL was prepared by mixing T4 DNA ligase (1 µL) and Ligase buffer (2 µL) to 50 ng restricted vector and suitable concentration of insert. Water was added to make it 20 µL and then it was incubated for overnight at 22°C.

3.3.6.1 Transformation of ligation reaction in E. coli

E. coli strain DH5α was employed for cloning purpose. Ligation reactions (5 µL) were transformed in competent cells of DH5α using heat shock. For transformation, 50 µL of competent cells were mixed with ligation mixture and left on ice for 30 minutes and then heat shock was applied at 42 °C for 60 seconds. Afterwards LB medium (150 µL) was poured to the transformed cells and further incubation was carried out at 37°C for 1 hour with shaking. LB plates supplemented with ampicillin (100mg/mL) were prepared and transformation mixture was spread on them followed by plates incubation at 37°C for overnight.

3.3.6.2 Confirmation of the positive clones

For the confirmation of positive clones, colony PCR was performed. A single colony was taken with the help of disinfected toothpick and agitated in 50 µl dH2O. All tubes were heated (5 minutes at 95°C) and spin down at 14000 X g for 5 minutes. The resultant supernatant was then employed as template and colony PCR
was performed using specific primers.

Positive colonies were transferred in LB media and were kept for overnight shaking at 37°C. Next day, plasmid was extracted using Plasmid isolation kit (Thermoscientific). Restriction digestion was carried out with BamHI and XhoI for 1 hour at 37°C to screen positive clones.

Insert and ORF were confirmed using pGEX-forward primer by commercial sequencing (Eurofins). Online BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and EXPASY translate tool (http://web.expasy.org/translate/) was used for sequence analysis.

3.4 OVER-EXPRESSION OF GST, GST-OsWRKY71 (167-271), GST-OsDOF18 (1-277), GST-OsDOF18 (1-148) and GST-OsMYB1 (88-222)

3.4.1 Transformation of Recombinant vectors in Expression Cells of *E. coli*

In order to choose the best strain to achieve high intensity of heterologous Protein expression, expression experiments were carried out by transformation of constructs into BL21 (DE3), BL21 (DE3) pLysS cells and BL21-CodonPlus (DE3) cells. One µL of recombinant plasmids were mixed with competent cells (50 µL) followed by incubation on ice for half an hour. LB plates supplemented with appropriate antibiotics i.e ampicillin (100 mg/mL) for all three types of cells used and Chloramphenicol (50 mg/mL) for BL21 (DE3) pLysS cells and BL21-CodonPlus (DE3) cells, were prepared. All transformed cells were subjected to spreading on antibiotics containing plates. For control Vector (native pGEX-4T1) transformation, BL21 (DE3) cells were used to express the GST protein, which was used in EMSA experiments as a negative control.
3.4.2 Expression of GST, GST-OsWRKY71, GST-OsDOF18 (1-277), GST-OsDOF18 (1-148) and GST-OsMYB1

A single colony from each plate was resuspended in LB medium having required antibiotics and cells were allowed to grow over night at 37°C (continuous shaking). In the growth medium 50 µM zinc chloride was added for the zinc-finger transcription factors, OsWRKY71 and OsDOF18. Next day, overnight culture (500 µL) was transferred to 50 mL of LB broth containing antibiotics and incubated at 37°C until OD$_{600}$ of 0.6-0.8 achieved. Recombinant proteins expression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to bacterial cultures and incubated at 37°C for further 3-4 hours at 250 rpm shaking. O.D$_{600}$ was checked after incubation time and cells were spin down (5000 X g for 20 minutes) at 4°C for harvesting purpose. Supernatant was discarded and pellets were stored at -80°C till future use.

Expression of GST-OsWRKY71 (167-271), GST-OsMYB1 (88-222), GST-OsDOF18 (1-277) and GST-OsDOF18 (1-148) was checked in all three kinds of E. coli cells. Samples (uninduced and induced) were checked by resolving on 15% SDS-PAGE (Appendix 2). In brief, 15% resolving gel and 4.5% stacking gels were permitted to polymerize for 2 hours. 2X SDS sample buffer was added to equalize samples concentration. Samples were prepared by heating (95°C for 5 minutes), followed by centrifugation. Samples were loaded on the SDS gel and molecular mass standard (GE Healthcare) was used to estimate the sizes of proteins. Electrophoresis was first performed at 60 V for 30 minutes and then voltage was increased upto 150 V for additional 45 minutes in Tris glycine SDS buffer (1X) (appendix 3). For staining purpose, gels were kept for 30 minutes in Coomassie Brilliant Blue R-250 (0.1% Coomassie R250; 10% acetic acid; 40% methanol) with gentle agitation and then into destaining solution (20% methanol; 10% acetic acid,
40 % methanol) until the background had vanished.

3.5 PURIFICATION OF GST, GST-OsWRKY71 (167-271), GST-OsDOF18 (1-148) and GST-OsMYB1 (88-222)

3.5.1 Cell Lysis

Cell pellet was thawed on ice and resuspended by adding 1X PBS (5 mL) supplemented with 1% Triton X-100. Sonication was performed to separate the recombinant OsWRKY71, OsDOF18 and OsMYB1. Sonication was carried out on ice 6 times for 30 seconds each with 30 seconds rest at maximum setting to lessen cells from heating. Therefore, samples were not sticky any longer. Then cells were spin down (20,000 Xg for 20 minutes) at 4°C. The resultant supernatant was saved at -80°C. PBS (1X) was added to the pellet to confirm the status of recombinant proteins on SDS PAGE.

3.5.2 Glutathione Sepharose Column Chromatography

Expressed GST, GST-OsWRKY71 (167-271), GST-OsDOF18 (1-148) and GST-OsMYB1 (88-222) were purified using Glutathione Sepharose 4B (GE Healthcare). Two mL bed volume of sepharose was used for 100 mL of culture. To prepare column, 2 mL of Sepharose was added to column and permitted it to straighten out. Extraction / wash buffer (20 mL) (phosphate buffer (50 mM), NaCl (300 mM) pH 7) was added to the column for equilibration of sepharose column. To the column, sonicated soluble supernatant was applied gradually, mixed well, poured to a falcon tube and left for shaking at 4°C. After 30 minutes, the mixture was transferred back onto the column and washing was carried out with wash buffer twice. After washing, recombinant protein was eluted in glutathione elution buffer (Appendix 4) in fractions of 1 mL. The column was washed three times with high
pH buffer (Appendix 5) and low pH buffer (Appendix 6) using 3 bed volume followed by re-equilibration with 1X PBS. Then column was stored at 4°C in 20% ethanol for reuse.

### 3.5.3 SDS-PAGE

The samples from each fraction were applied to 15% SDS PAGE to check purity. Fifty µL of SDS sample buffer (5X) was combined with all samples before being heated at 95°C for 5 minutes. Molecular mass standard (GE Healthcare) was used to estimate the sizes of proteins. Electrophoresis was performed at 60 V first for 30 minutes and then the voltage was increased up to 150 for additional 45 minutes in 1X Tris glycine SDS buffer. Gels were stained on a rocking platform for 30 minutes with Coomassie Brilliant Blue R-250 (0.1% Coomassie R 250; 10% acetic acid; 40% methanol aqueous solution) and then into the destaining solution (20% methanol; 10% acetic acid aqueous solution) until the background had disappeared.

### 3.5.4 Cationic Exchange Chromatography

GST-OsWRKY71 (167-271) and GST-OsMYB1 (88-222) were further purified by cation exchange chromatography. GST-OsWRKY71 (167-271) and GST-OsMYB1 (88-222) elutions were dialyzed for 24 hours in phosphate buffer (phosphate buffer (50 mM) pH-7, NaCl (0.1 M)). Dialyzed proteins were subjected to centrifugation at 14000 X g for 10 minutes to remove dust particles if any. SOURCE\textsuperscript{TM} 15S column (GE Healthcare) packed with Polystyrene/divinyl benzene was used. The column was first equilibrated with Phosphate buffer. Then proteins were injected to SOURCE\textsuperscript{TM} 15S column followed by washing with phosphate buffer. Linear NaCl gradient of the buffer was applied to elute out the bound proteins up to 1M concentration. Fractions of eluted proteins were checked on 15% SDS PAGE for purity. Purified fractions were pooled and saved at -20°C.
3.5.5 Purified GST, GST-OsWRKY71 (167-271), GST-OsMYB1 (88-222), and GST-OsDOF18 (1-148) Dialysis and Concentration Determination

Purified proteins were subjected to dialysis in EMSA binding buffers. Purified GST, GST-OsWRKY71 (167-271) and GST-OsDOF18 (1-148) were dialyzed in buffer having composition of HEPES-KOH (25 mM) pH 7.2, KCl (40 mM), EDTA (1 mM), DTT (1 mM), ZnCl$_2$ (50 µM) and Glycerol (10%). GST and GST-OsMYB1 (88-222) were dialyzed in HEPES (25 mM) pH 7.2, KCl (4 mM), MgCl$_2$ (5 mM), EDTA (1 mM), DTT (0.5 mM) and glycerol (7%). Standard RC Dry Dialysis Tubing of 6-8 kDa MWCO was used. Dialysis was performed for 12 hours at 4°C in such a way to change buffer twice after 4 hours and then left it overnight in cold room.

After dialysis, absorbance was recorded at 280 nm using nanodrop to estimate the concentration of the dialyzed purified proteins with extinction coefficients envisaged by Protparam online tool (http://web.expasy.org/protparam/). Proteins were subjected to be concentrated using 3 K AMICON filters at 14000 g each round for 10 minutes and at 4°C.

3.6 PURIFIED PROTEINS CONFIRMATION BY MASS SPECTROMETRY

3.6.1 On Column Cleavage of GST Tagged Proteins

All recombinant proteins expression and purification was carried out as described previously till the mixing and incubation of sonicated supernatant with sepharose beads. After incubation, mixture was washed with Thombin cleavage
buffer (1X PBS) and then buffer was eluted out till 5 mL left in column. Washed mixture was transferred to a falcon tube. Thrombin solution (80 units) by mixing 50 µL of thrombin with 950 µL of thrombin cleavage buffer. Falcon tube was incubated at room temperature for 4 hours, followed by mixture transfer on to column and collection of flowthrough. Column was washed with 1X PBS and run through was collected. Glutathione bound to column was then eluted with reduced glutathione buffer. To remove the rest over of thrombin from cleaved proteins, benzamidine column was used. First benzimidine sepharose was added to column and allowed it to settle down. Sepharose beads were equilibrated with PBS buffer and eluants were loaded onto column. Flow through was collected in fractions and all fractions were checked by running on 20 % SDS PAGE.

3.6.2 Analysis on Mass Spectrometer

3.6.2.1 Target cleaning and matrix preparation

Target plate was first cleaned with water and paper towel followed by washing with ethanol (96%). Matrix was prepared by mixing 1 volume of 100 acetonitril with 1 volume of 1% Trifluoroacetic acid. Two µL of cleaved proteins were mixed with 2 µL sinapinic acidfollowed by mixing with 0.1% trifluoroacetic acid on a parafilm. Then 2µL of matrix was mixed with 2µL protein sample and loaded on the target plate. Position of all samples on target plates was noted. Drops were air dried and identification of each protein was carried out by MALDI-TOF analysis.

3.6.2.2 MALDI TOF Mass Spectrometry

Bruker REFLEX III MALDI mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) was used to analyze the cleaved purified proteins. The
fragmentation data were analyzed using a standard peptide mixture (Sigma-Aldrich, St. Louis, MO, USA) was used to calibrate MALDI-TOF spectra. This standard mixture includes insulin (m/z 5734.6), adrenocorticotropic hormone fragment 18–39 (m/z 2466.7) and angiotensin II (m/z 1047.2). After calibration, BioTools version 2.2 software (Bruker Daltonics) was used to interpret all spectra.

3.7 ELECTROPHORETIC MOBILITY SHIFT ASSAYS (EMSAs)

Purified recombinant OsWRKY71 (167-271), OsDOF18 (1-148) and OsMYB1 (88-222) proteins were employed in Electrophoretic mobility shift assays (EMSAs). DIG gel shift kit (2nd generation) was used for this purpose with different alterations that are mentioned below with each step.

3.7.1 Probe Designing

Regulatory cis-elements containing short fragments of OsRGLP2 promoter of about 30bp that form binding site for specific transcription factors were designed. Detail of all these fragments is mentioned in table 2. Three probes were designed for OsWRKY71; one including both true W-box (C/TTGACT/C) and core W-box (TGAC) while other two probes contain either true W-box (C/TTGACT/C) or core W-box (TGAC) sequences. These fragments were commercially synthesized in HPLC purified form. For OsDOF18, two probes were designed; one containing two AAAG motifs while other containing only one AAAG motif. Two probes were designed for OsMYB1; one containing TAACCA binding motif while other contains AAACCA binding motif. In all cases, core cis-elements were mutated in probes to check the specificity of DNA-protein interactions. Sequences of mutated probes are also mentioned in table 2.
3.7.2 Probes Labelling and Purification

Commercially synthesized probes were in lyophilized form. TEN buffer (Appendix 7) was added to these probes to dissolve them in a final concentration of 100 pmol/µL. Equal moles of single stranded oligonucleotides of complementary sequences were mixed and heated at 95°C in a water bath for 10 minutes. This mix was then annealed by allowing them to cool slowly to room temperature left at room temperature for 6 hours. Annealed probes were then stored at -20°C.

For probes labeling, 100 pmol of annealed probes were mixed with 4 µL of labeling buffer (5X), 4 µL of CoCl (25mM), 1 µL of DIG-11-ddUTP (1mM) and 1 µL of Terminal Transferase enzyme in a total reaction volume of 20 µL. This labeling reaction mix was kept for 1 hour at 37 ºC followed by blocking the reaction by adding 2 µL of EDTA (0.2 M, pH 8).

For labeled probes purification, method of ethanol precipitation was used. For this purpose, labeled probes were mixed with 2 µL of LiCl (4M) and 60 µL of 100% ice-cold ethanol (100%). The reaction was incubated for 2 hours at -80°C to facilitate precipitation. After incubation, the reaction was spin down at 13000 X g for 15 minutes at 4°C to pellet down the precipitated DNA. 100 µL of ethanol (70 %) was added to wash the pellet of labeled precipitated probes followed by centrifugation again for 15 minutes at 13000 X g at 4 ºC. The pellet was then allowed to air dry completely and dissolved in TEN buffer to make it 2.5 pmol/µL (final concentration).

3.7.3 Labelling Efficiency Determination

Newly labeled probes were checked for their efficiency by dot blot. For this purpose, different dilution series of labeled probes were made and spotted on the
Table 2: *OsRGLP2* promoter fragments used as probes in Electrophoretic Mobility Shift Assays (EMSAs)

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Sequence</th>
<th>Position on promoter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><em>OsWRKY71-F1</em></td>
<td>5'-GTGTGACCAAGTAAATTTTTGACTAGTCTG-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsWRKY71-R1</em></td>
<td>5'-CAGACTAGTCAAAAAATTATTTGTCACAC-3'</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td><em>OsWRKY71-F2</em></td>
<td>5'-CATATGCTACCTGACCGTACTGGTGACGA-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsWRKY71-R2</em></td>
<td>5'-TCCGTACCAGTACGGTGACGATGACATATG-3'</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td><em>OsWRKY71-F3</em></td>
<td>5'-CTTTTCGATCGAGTACCGTAAAGGCTGCTG-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsWRKY71-R3</em></td>
<td>5'-TGACCTGCATGCTAAGTGATCGATCGAAAG-3'</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td><em>OsDOF-F1</em></td>
<td>5'-GATATGCTAAAGTGACAAAGCTGCTGCTG-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsDOF-R1</em></td>
<td>5'-CGACGAGCTTTTGCTACCTTTTAGCATATC-3'</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td><em>OsDOF-F2</em></td>
<td>5'-ATTGCGAGTGAAGTTATGAAACCACTCCCC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsDOF-R2</em></td>
<td>5'-GGGAGTTGTTAATTCTCTTCGAGAT-3'</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td><em>OsMYB1-F1</em></td>
<td>5'-GACATAATTGGAACCTGACTGTTAAACT-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsMYB1-R1</em></td>
<td>5'-AGTTTAACTGACTGTTTTCCCATATGTC-3'</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td><em>OsMYB1-F2</em></td>
<td>5'-CCAGTGAAGTAAATTTACCTCCAGAAG-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsMYB1-R2</em></td>
<td>5'-TTCTGGGAGTGTAGTTAATTCTTTCACTG-3'</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td><em>OsWRKY71-F1m</em></td>
<td>5'-GTGAAAAAAATGAAATTTTTTAAAAAGTCTG-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsWRKY71-R1m</em></td>
<td>5'-CAGACTTTTTAAAAATTTTAAATTTTACCT-3'</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td><em>OsDOF-F1m</em></td>
<td>5'-GATATGCTAGAGGACGATAGAGCTGCTG-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsDOF-R1m</em></td>
<td>5'-CAGACGAGCTTTGCTACTGCTTTAGCAATC-3'</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td><em>OsMYB1-F1m</em></td>
<td>5'-GACATAATTGGAACCTGACTGTTAAACT-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>OsMYB1-R1m</em></td>
<td>5'-AGTAAAACTGACTGTTTTCCCATATGTC-3'</td>
<td></td>
</tr>
</tbody>
</table>
positively charged membrane. Comparison was made with one labeled control provided with the kit. 1µL of each serial dilution was spotted to small piece of positively charged nylon membrane followed by fixing of probes to the membrane by crosslinking in the presence of UV light. Nylon membrane strip was then washed with washing buffer (20 mL) (Appendix 8) in a falcon tube with constant agitation at room temperature for 2 minutes. After washing, blocking solution (1X) (Appendix 9) was added to the membrane and left on shaking for 30 minutes at room temperature. After blocking, 10 mL of antibody solution was added and again incubated at room temperature for 30 minutes. The membrane was thoroughly washed two times with washing buffer (50 mL) for 20 minutes each time followed by addition of detection buffer (Appendix 10) for equilibration. A plastic folder was cut and kept in a tray. Membrane was transferred to a plastic folder after equilibration and CSPD solution was added onto membrane slowly followed by squeezing out of excess of liquid. Then to enhance the Chemiluminescent, membrane was incubated at 37°C. After 20 minutes incubation, membrane was exposed to X-ray film for 3 hours. To develop film, it was dipped in developer solution first for few seconds, then in fixer solution again for few seconds followed by washing by dipping in water. This whole procedure of developing film was done in a dark room with red light only.

3.7.4 Binding Reaction Preparation and Optimization

To optimize the binding of labeled probes with recombinant purified proteins, various binding buffers were used to find the most appropriate buffer. The recipes of all buffers screened in this study are mentioned in table 3. To reduce the chances of pipetting error, a master mix was prepared by mixing all common components of same concentration and then required amount of this mixture was added to each reaction tube. The binding mixture containing 1 μg of recombinant protein was incubated on ice with 4 μL of binding buffer (5X), 1 μg of
Table 3: Buffers screened for EMSA

<table>
<thead>
<tr>
<th>No.</th>
<th>GST-OsWRKY71, GST-OsDOF18</th>
<th>GST-OsMYB1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25 mM HEPES-KOH pH 7.2, 40 mM KCl, 1 mM DTT, 1 mM EDTA, 10% Glycerol (1X)</td>
<td>50 mM Tris-HCl pH 7.5, 40 mM MgCl₂, 5 mM EDTA, 5 mM DTT, 250 mM NaCl, 25% glycerol (5X)</td>
</tr>
<tr>
<td>2</td>
<td>50 mM Tris HCl pH 7.5, 250 mM NaCl, 2.5 mM DTT, 2.5 mM EDTA, 5 mM MgCl₂, 20% Glycerol (5X)</td>
<td>25 mM HEPES pH 7.2, 4 mM KCl, 5 mM MgCl₂, 1 mM EDTA, 7% Glycerol, 0.5 mM DTT (1X)</td>
</tr>
<tr>
<td>3</td>
<td>10 mM Tris HCl pH 8, 20 mM NaCl, 0.4 mM MgCl₂, 0.5 mM ZnSO₄, 0.25 mM EDTA, 0.25 mM DTT, 10% Glycerol (1X)</td>
<td>10 mM TrisHCl pH 8, 1 mM EDTA, 1 mM DTT, 50 mM NaCl, 5% Glycerol (1X)</td>
</tr>
<tr>
<td>4</td>
<td>100 mM HEPES pH 7.6, 5 mM EDTA, 50 mM (NH₄)₂SO₄, 5 mM DTT, 1% Tween 20, 150 mM KCl (5X)</td>
<td>100 mM HEPES pH 7.6, 5 mM EDTA, 50 mM (NH₄)₂SO₄, 5 mM DTT, 1% Tween 20, 150 mM KCl (5X)</td>
</tr>
</tbody>
</table>
the poly d(I-C), 0.1 µg of Poly L-lysine and 50 fmol of labeled probe. Then reaction mixture was subjected to pipetting up and down, spins down the reaction and kept at 25°C for 30 minutes.

For competition analysis, 200 molar excess of unlabeled probes were added to the reaction mixture followed by 15 minutes incubation on ice before the addition of labeled probe. After incubation, labeled probe was added to this reaction mixture followed by incubation of further 15 minutes at 25°C.

3.7.5 Native Polyacrylamide Gel Electrophoresis (Native PAGE)

To avoid any finger prints and dust particles, short and spacer plates of PAGE were washed thoroughly with ethanol (70%) and allowed to air dry. Acrylamide and bisacrylamide polymerization mix was made. For EMSA, 6% native gel was used, recipe of which is provided in appendix 11. Mixture was poured into assembly immediately after the addition of TEMED and APS (10%) followed by right away putting in comb and then gel was permitted to polymerize for minimum 2 hours. When the gel was polymerized, the comb was removed and wells were washed. The gel was prerun at 4°C for 30 minutes at the voltage of 80 V in 0.5 X TBE buffer (Tris (89 mM), borate (8.9 mM), EDTA (2 mM), pH 8). To stop the reaction of EMSA, loading buffer (3 µL) containing bromophenol blue was mixed in reaction and all EMSA reactions were loaded on prerun native gel instantly. Voltage (80 V) was applied to the gel to perform electrophoresis at 4°C for approximately 90 minutes.

3.7.6 Blotting and Crosslinking

Transblot semidry electrophoretic transfer cell from Bio-Rad was used for
Electro-blotting. Equilibration of 5 minutes was performed for positively charged nylon membrane and gel sized Whatman3 MM papers in transfer buffer i.e 0.5X TBE. Equilibrated whatman papers were placed on electrophoretic transfer cell followed by nylon membrane. The gel was transferred carefully onto a membrane and further layers of Whatman papers were added onto the gel. To keep away any bubbles, a pipette was rolled over the Whatman paper once. To transfer the gel on nylon membrane, a constant voltage of 30V was provided along with 300 mA current and 10 Watt for 30 minutes.

When gel transfer onto membrane was complete, membrane was placed onto a Whatman paper. This Whatman paper was presoaked in 2X saline-sodium citrate (SSC) buffer (Appendix 12). To crosslink the DNA to membrane, it was exposed to UV light for crosslinking at 1200 µjoules in a UV stratalinker.

### 3.7.7 Chemiluminescent Signal Detection and X-ray Exposure

About 30 mL of washing buffer was added to the plate containing nylon membrane and incubated at room temperature for 5 minutes for proper washing. Then washed membrane was dipped in blocking solution (2X) and left at room temperature with constant agitation for 1 hour. When blocking was complete, a solution of Alkaline Phophatase conjugated Anti-Digoxigenin antibody (diluted 1:20,000 in blocking solution) was added onto membrane. After 30 minutes incubation with antibody solution, membrane was washed two times with washing buffer (100 mL each time) for 20 minutes with gentle shaking. Then detection buffer (50 mL) was added for 5 minutes to equilibrate the membrane. A plastic folder was cut and was placed in developing cassette. Membrane was transferred carefully onto plastic folder and chemiluminescent AP substrate (CSPD solution
diluted 1:100 with detection buffer) was pipette drop wise (DNA side up) in such a way to cover the edges first. Then the entire membrane surface was covered with this solution by tilting the membrane. Excess substrate was wiped out using a paper towel to reduce background. Exposure Cassette along with membrane was kept for 5 minutes at room temperature followed by sealing of the edges of developing folder having membrane with DNA side up. After that, the membrane was left at at 37°C for 15-30 minutes to enhance the signals. Then membrane was overlayed with an X-ray film (Hyperfilm-MP, Amersham) in a darkroom with the safe light on. Exposure was allowed for 3 hours. Then the X-ray film was developed by dipping the film first in developer solution followed by transfer in fixer solution for short time. Then film was washed with water.

3.8 EMSA WITH NUCLEAR PROTEINS EXTRACT

3.8.1 Nuclear Proteins Extraction

To isolate nuclear proteins extract from rice, 10 days old rice seedlings were used. First these seedlings (5g) were crushed in precooled mortar and pistil with the help of liquid nitrogen. This homogenized material was then transferred to a falcon tube containing extraction buffer (15 mL) (Appendix 13) supplemented with PMSF (0.2 mM) and mixture of protease inhibitor (1 mg/mL). This homogenized solution was filtered through 2 layers of miracloth tissue of 100-mm pore size. Flowthrough was transferred to a falcon tube and spin for 10 minutes at 4300 X g (4°C) in a precooled rotor. Supernatant was removed carefully and pellet was resuspended in protein isolation buffer (500 µL) (Appendix 14) supplemented with PMSF (0.2 mM) and mixture of protease inhibitor (1 mg/mL). Nuclei suspension was incubated for 40 minutes at 4°C. After every 6 minutes falcon tube
was stirred for few seconds using vortex. Then nuclei solution was shifted to Eppendorf tubes (1.5 mL capacity) and centrifugation was performed for 15 minutes at 12,000 X g(4°C) in a precool rotor. After transferring supernatant to a new tube, glycerol was added up to 20% final concentration. Nuclear extract was then aliquoted into small fractions (100 µL) and placed at -80°C till further use.

For EMSA reactions, dialysis of nuclear proteins was performed in EMSA binding buffer (HEPES-KOH (25 mM) pH 7.2, KCl (40 mM), DTT (1 mM), EDTA (1 mM), 50 µM ZnCl₂, Glycerol(10%) for 24 hours at 4°C. Standard RC Dry Dialysis Tubing of 6-8 kDa MWCO was used. Dialysis was performed for 12 hours at 4°C in such a way to change buffer twice after 4 hours and then left it overnight in cold room.

Absorbance was recorded at 280 nm using nanodrop to estimate the concentration of the dialyzed nuclear proteins extract.

3.8.2 Binding Reaction and Detection

Twenty µL EMSA reaction was prepared by mixing 7 µg of nuclear proteins extract with 4 µL of binding buffer (5X), poly d(I-C) (1 µg), Poly L-lysine (0.1 µg) and labeled probe (50 fmol). Reaction was mixed by pipetting up and down, spinning down the reaction. After spin, reaction was kept for 30 minutes at 25°C (room temperature).

For competition analysis, EMSA reactions were prepared in the same way as illustrated previously with the exemption that the nuclear proteins extract was kept at room temperature for 15 minutes with the unlabeled Oligonucleotides (200
molar excess) before the addition of labeled probe to the reaction. Reactions were loaded on 6% native PAGE and similar procedure was repeated as described in section 3.6.

3.9 TARGET GENES EXPRESSION PROFILING

Two approaches were used to investigate the expression level of OsWRKY71, OsDOF18 and OsMYB1 genes. One approach was to get data from GENEVESTIGATOR tool (https://genevestigator.com/gv/plant.jsp). All the data in GENEVESTIGATOR databases have been obtained from public microarray repositories and it contains data from many organisms including rice. Second approach was to validate the microarray expression patterns of chosen genes under different abiotic stresses by quantitative real-time PCR (qRT-PCR).

3.9.1 Microarray Data Retrieval from GENEVESTIGATOR

Microarray data for selected genes was retrieved from the microarray database known as GENEVESTIGATOR (Plant Biology) (https://genevestigator.com/gv/plant.jsp). OsWRKY71, OsDOF18 and OsMYB1 found to be up regulated under various stress conditions having probe sets of Os02g0181300, Os08g0490100 and Os01g0850400 respectively. Only abiotic stresses related data was chosen for instance, cold, drought, heat and salinity.

3.9.2 Quantitative Real-Time PCR (qRT-PCR)

3.9.2.1 Primer designing

The gene specific primers for OsWRKY71, OsDOF18 and OsMYB1 were designed to amplify specific regions. Actin was used as endogenous control in qRT-PCR and one set of primer was designed for it also.
Table 4: Primers designed for Real Time PCR

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Length</th>
<th>Tm</th>
<th>GC (°C)</th>
<th>Product size</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actin-F</td>
<td>GAAGATCACTGCCTTGCTCC</td>
<td>20</td>
<td>55.9</td>
<td>55</td>
<td></td>
<td>X16280.1</td>
</tr>
<tr>
<td>Actin-R</td>
<td>CGATAACAGCTCCTCTTGCC</td>
<td>20</td>
<td>55.6</td>
<td>55</td>
<td>226</td>
<td></td>
</tr>
<tr>
<td>WRKY71-F</td>
<td>TGGATTAGCACCCAGCCTTC</td>
<td>20</td>
<td>57.3</td>
<td>55</td>
<td></td>
<td>AB190817</td>
</tr>
<tr>
<td>WRKY71-R</td>
<td>AGGCTGCTGGTGAAAGAAGT</td>
<td>20</td>
<td>56.8</td>
<td>50</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>MYB1-F</td>
<td>CGAGGAACTGGACGCTGATA</td>
<td>20</td>
<td>56.4</td>
<td>55</td>
<td>152</td>
<td>NM_001051349</td>
</tr>
<tr>
<td>MYB1-R</td>
<td>ATGGATTGCATGACGACCGCA</td>
<td>20</td>
<td>57.6</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOF18-F</td>
<td>AAGACGACGACTCCTCCACAAC</td>
<td>20</td>
<td>55</td>
<td>50</td>
<td>182</td>
<td>NM_001068645</td>
</tr>
<tr>
<td>DOF18-R</td>
<td>AGACTCTTGGTGGATGACCGG</td>
<td>20</td>
<td>56.5</td>
<td>55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
their properties were performed using different bioinformatics tools i.e. Integrated DNA technology oligoanalyzer (www.idtdna.com) or Primer BLAST (www.ncbi.nlm.nih.gov/primerBLAST). Sequences spanning the two exons junction were chosen to enhance specificity. Sequences and information about all primers used in RT-qPCR is provided in table 4.

3.9.2.2 Plant growth conditions and stress treatments

*Oryza sativa* cv. KS282 seeds were grown in test tubes on half strength MS basal medium (Murashige and skooge, 1962) for 10 days in growth room at 25±1°C. Then various abiotic stresses were applied on 10 days old rice seedlings, detail is given below.

For salt stress, 10 days old seedlings were transferred to a beaker contain NaCl (200 mM) solution and were kept in this condition at 25± 1°C for 3 hours. After salt treatment, samples were collected from control and treated seedlings for RNA extraction, weighing was done, and placed at -80°C immediately after exposing to liquid nitrogen.

For wounding stress, 7 mm pieces of leaves were cut from 10 days old seedlings, kept in a water containing beaker, and allowed to float under suitable light conditions for a period of 6 hours at room temperature. Samples were collected from control and treated seedlings for RNA extraction, weighing was performed and placed at -80°C immediately after exposing to liquid nitrogen.

Ten day old seedlings were exposed to 4°C for cold stress for the period of 48 hours. Then rice seedlings were transferred to room temperature for 24 hours.
Samples were collected from control and treated seedlings for RNA extraction, weighing was done, and placed at -80°C immediately after dipping in liquid nitrogen.

For heat stress, 10 day old seedlings were transferred to 45°C for the duration of 6 hours. samples were collected from control and treated seedlings for RNA extraction, weighing was performed and placed at -80°C immediately after exposing to liquid nitrogen.

Drought stress was applied to 10 day old seedlings by withholding water supply and placing the seedlings onto a paper towel till observable leaf rolling emerged in the plants. samples were collected from control and treated seedlings for RNA extraction, weighing was performed and kept at -80°C immediately after exposing to liquid nitrogen.

3.9.2.3 RNA Extraction and determination of RNA quality

Total RNA was extracted from 10 days old rice seedlings using the RNeasy mini kit from QIAGEN according to the instructions provided with certain modifications as described in section 3.3.3. 1.2% agarose gel (prepared with RNAse free 1X TAE buffer) was run with isolated RNA samples to examine the quality and integrity of RNA. Loading dye was mixed with RNA samples, heated at 70°C for 3 minutes and kept on ice immediately. Electrophoresis was performed at 66V for 30 minutes. After gel running, it was exposed to UV for band visualization. Total RNA concentration was measured using nanodrop by taking absorbance at 260 nm.

3.9.2.4 Real time PCR

Brilliant II SYBR® Green QRT-PCR Master Mix Kit was used to perform
quantitative RT-PCR. RT-PCR reactions consisted of QPCR mastermix (5 µL), RNA (100 µg), forward and reverse primers (0.5 µM each) and RT/RNase block enzyme mixture (0.2 µL) in 20 µL reaction volume. Amplification procedure consisted of synthesizing cDNA by pre-incubating the reaction mixture for 30 minutes at 50°C, then 10 minutes at 94°C, followed by 40 cycles of 30 seconds 94°C (Denaturation), 1 minute 53°C (Annealing), and 30 seconds 72°C (extension). Melting curves were obtained in the range of 65-95°C as last step of RT PCR to check the specificity of amplification. CFX96 real-time PCR detection system (Bio-Rad) was used to perform all reactions. Internal control was a house keeping gene i.e β-actin. To evaluate the repeatability of RT-PCR, assay was performed three times, with sample three replicates for each sample.

### 3.9.2.5 Data analysis of qRT-PCR

The 2^{−ΔΔCT} method was employed to interpret the qRT-PCR data according to the method of Schmittgen and Livak, (2008). Expression data of OsWRKY71, OsDOF18 and OsMYB1 were standardized by taking away the mean of CT value of reference gene from mean of individual CT values of the consequent gene of interest (ΔCT). The value of fold-change was intended using the 2^{−ΔΔCT} where ΔΔCT represents the difference between the ΔCT condition of gene of interest and ΔCT of internal control.

### 3.10 ANALYSIS OF E. COLI TRANSFORMANTS IN ABIOTIC STRESSES

The role of OsWRKY71, OsMYB1 and OsDOF18 in E. coli cells was examined by spot assay. Recombinant plasmids including pGEX4T1-OsWRKY71, pGEX4T1-OsDOF18 and pGEX4T1-OsMYB1 and control plasmid i.e pGEX4T-1 were transformed in BL21 (DE3) cells. Cells were grown in LB broth till OD_{600} 0.6
was attained. Afterward, 1 mM IPTG was added to all cultures to induce the expression of recombinant proteins and incubation was extended for more 4 hours at 37°C. Bacterial cultures were diluted to OD_{600} 1 after measuring the absorbance of all cultures at 600 nm. Then 50-fold, 100-fold and 200-fold dilutions were made for all bacterial cultures.

LB plates were made supplemented with ampicillin (50 µg/mL) and IPTG (1 mM) in addition to 400, 500, and 600 mM NaCl concentration gradient for salt stress analysis. All dilutions (10 µL) were then pipetted as spot on IPTG LB agar plates followed by incubation at 37°C for overnight.

LB IPTG plates were supplemented with 500 mM, 800 mM, and 1M mannitol concentration gradient to examine the drought stress of transformants. All dilutions (10 µL) were then pipetted as spot on IPTG LB agar plates followed by incubation at 37°C for overnight.

Multiple aliquots of 1mL were placed at 50°C for heat stress. After 1, 2 and 3 hours of incubation, 100 µL sample from each aliquot was taken for further analysis. 50-fold, 100-fold and 200-fold dilutions were prepared for all samples and all dilutions (10 µL) were then pipetted as spot on IPTG LB agar plates followed by incubation at 37°C for overnight.

Samples were kept at -80°C for cold stress for the duration of 24 hours. Then all samples were taken out and kept at 35°C for 60 minutes followed by taking aliquots of 100µL for each sample after 2, 4, 6 and 8 hours successively. 50-fold, 100-fold and 200-fold dilutions were prepared for all samples. All dilutions (10 µL)
were then pipetted as spot on IPTG LB agar plates followed by incubation at 37°C for overnight.

3.11 **IN SILICO CHARACTERIZATION OF IDENTIFIED PROTEINS**

Primary sequences of *OsWRKY71*, *OsDOF18* and *OsMYB1* were analyzed by DISOPRED ([http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1](http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1)) to measure the degree of intrinsic disorder. Protein-protein interaction networks were constructed by STRING ([http://string-db.org/](http://string-db.org/)). The interactions include physical and functional associations.
RESULTS AND DISCUSSION

Plant growth and yield are significantly influenced by stresses for instance wounding, drought, salinity, cold and pathogen infection. To minimize damage caused by these harmful factors, plants respond by reprogramming the expression level of stress related genes via various transcription factors. In recent times, the functions of a growing number of stress responsive genes and transcription factors are being revealed. Understanding of these mechanisms is vital for the progress of transgenic approaches to increase the stress tolerance in crop plants. Germin like proteins (GLPs) are associated with different biotic and abiotic stresses. The promoter of the OsRGLP2 gene has revealed strong GUS expression under different abiotic stresses. In order to get functional insight, the present research was undertaken to identify, clone, express and characterize the stress responsive OsRGLP2 promoter binding proteins.

4.1 IN SILICO ANALYSIS OF OsRGLP2 PROMOTER FOR STRESS RESPONSIVE REGULATORY MOTIFS AND CORRRESPONDING BINDING PROTEINS

In silico analysis has been broadly used for the identification of putative regulatory cis-elements in the promoter section of a gene (Ibraheem et al., 2010). Several functional putative cis-acting regulatory elements for stress related transcription factors were recognized in the OsRGLP2 promoter by using PLACE, PlantCARE, Consite and JASPAR online databases. All types of stress responsive cis-acting elements and their role in different stresses is described in table 5. W-box, WAACCA and AAAG motifs were observed to be prevalent at varying frequencies on positive and negative strands of the OsRGLP2 promoter region (Figure 1).
Table 5: Stress responsive cis-acting elements and their corresponding binding factors located on OsRGLP2 promoter and their role in different stresses

<table>
<thead>
<tr>
<th>Name</th>
<th>Cis element</th>
<th>Factor</th>
<th>Family</th>
<th>Stress signal</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>W-Box</td>
<td>C/TTGACC/T</td>
<td>WRKY</td>
<td>WRKY</td>
<td>Wounding, pathogen infection</td>
<td>Hara et al., 2000; Rushton et al., 1996</td>
</tr>
<tr>
<td>MYB</td>
<td>WAACCA</td>
<td>MYB</td>
<td>MYB</td>
<td>Involved in regulation of drought inducible gene expression</td>
<td>Seo et al., 2009</td>
</tr>
<tr>
<td>DRE</td>
<td>RCCGAC</td>
<td>DRE</td>
<td>DRE</td>
<td>Function in drought-, high-salt- and cold-responsive</td>
<td>Hong et al., 2006</td>
</tr>
<tr>
<td>ABRE</td>
<td>ACGTG</td>
<td>ABRE</td>
<td>bZIP</td>
<td>ABA responsive</td>
<td>Lopez-Molina et al., 2001</td>
</tr>
<tr>
<td>DOFZM</td>
<td>AAAG</td>
<td>DOF</td>
<td>Zinc finger</td>
<td>Development, drought and salt stress</td>
<td>Kang et al., 2003</td>
</tr>
</tbody>
</table>
Figure 1: Position of different stress responsive cis-elements on 1107 bp OsRGLP2 promoter region
Fourteen putative W-boxes, fourteen AAAG and three WAACCA motifs were found in OsRGLP2 promoter. W-boxes with typical TGAC core motif, form binding site for WRKY proteins, are reported to be involved in plant development, abiotic and biotic stresses (Eulgem and Somssich, 2007; Himmelbach et al., 2010).

The AAAG motif forms a core site for DOF proteins binding which belongs to a type of zinc finger regulatory protein family that perform various roles in plants such as in developmental stages (Konishi and Yanagisawa, 2007) and different stresses (Kang et al., 2003). MYB transcription factors bind to WAACCA motif and respond to various environmental stimuli (Gomez-Maldonado et al., 2004) and plant metabolism (Goicoechea et al., 2005). Therefore, it is rational to suppose that WRKY, DOF and MYB proteins are implicated in regulating the transcription of the OsRGLP2 gene under abiotic stresses.

4.2 RETRIEVAL AND SEQUENCE ANALYSIS OF OsWRKY71, OsDOF18 AND OsMYB1 TRANSCRIPTION FACTORS

Sequences of proteins that putatively bind with W-box, MYB box and AAAG motifs were retrieved from NCBI and the structure of DNA binding domain was analyzed. DNA binding domains recognize the regulatory motif sequence and bind with it through short motif, mostly an α-helix or a β-sheet. These short motifs make contact with the major groove of double-stranded DNA. Usually the DNA-protein contact is spread out across 5 bp, with a relatively high affinity and sequence-specificity.

4.2.1 Sequence Analysis of OsWRKY71

OsWRKY71 was predicted by PLACE as an interaction partner of W-box in the OsRGLP2 promoter. To better understand about the WRKY 71 DNA binding domain, multiple sequence alignment was carried out with WRKY proteins from
different plants and it was observed that the DNA binding domain sequence is highly conserved in all WRKY proteins. Secondary structure was predicted by using PSIPRED and revealed that the DNA binding domain of OsWRKY71 is made up of ~60 amino acid residues extending from Valine191 to Proline252 and comprises of four β-sheets. It includes a single, highly conserved WRKY domain at the N terminus and zinc finger like structure at its C terminus (C-X₅-C-X₂₃-H-X₁-H) which means that it belongs to group IIA (Eulgem et al., 2000) (Figure 2).

The heptapeptide WRKYGQK forms the first β-strand and is involved in making contact with major groove of DNA. C218, C224, H248 and H250 are involved in binding with Zn²⁺ and consequently forming the Zn finger motif. Maeo et al. (2001) observed that WRKY DNA binding domain has abolished its DNA binding activity when conserved Cys and His residues were mutated in Zinc finger motif. Basic residues K204, R213, R217 and R227 make contact with the backbone. Zhang et al. (2004) functionally characterized OsWRKY71 and found to be overexpressed in aleurone cells and act as a negative regulator of GA signaling. Liu et al. (2007) observed that OsWRKY71 expression was up-regulated in case of wounding and pathogen infection in rice.

4.2.2 Identification and Sequence Analysis of OsMYB1

Based upon the putative cis-elements binding sites, sequences of AtMYB1 (predicted by PLACE, Consite and JASPAR) was retrieved from NCBI database. In order to identify MYB protein encoded by the rice genome, the amino acid sequences of AtMYB1 was subjected to a BLAST survey with limitation to include hits only from Oryza sativa. The result showed AtMYB1 to have 42% sequence similarity in the DNA binding domain region with Os01g0850400 which is similar to partial OsMYB1 protein sequence (Accession no. CAE00856.1). OsMYB1 encodes a protein composed
Figure 2: Protein sequence alignment of OsWRKY71 structure of DNA binding domain

WRKY proteins from different plant members including WRKY from T. aestivum, H. vulgare, S. bicolor, B. distachyon and Z. mays were analyzed by TCOFFEE. Genbank accession numbers for OsWRKY71, TaWRKY80, HvWRKY1, SbputativeWRKY, TaWRKY8, BdWRKY40 and ZmWRKY71 are DAA05136.1, AFW98256, AAS48544.1, XP_002451666.1, ABC61128.1, XP_003570741.1 and NP_001147732.1 respectively. Conserved residues are shaded in different colors. Green arrows indicate the four β strands of DNA binding domain in the C terminus of OsWRKY71. Arrows represent the key residues involve in making contact with DNA major groove. Arrowheads denote residues that make contact with DNA backbone. Asterisks represent cysteine and histidine residues of Zinc finger like motifs in WRKY DNA binding domain.
of 413 amino acids with a predicted molecular weight of 44.3 kDa and estimated pI of 6.24.

Multiple sequence alignment with MYB proteins from different plants revealed that it contains two MYB DNA binding domains at N-terminus and belongs to R2R3-type subfamily of MYB DNA proteins (Figure 3). Both DNA binding domains (R2 and R3) are essential for binding to specific DNA sequences. Conserved and regularly spaced tryptophan residues in DNA binding domain play role in folding of domain by making a cluster of helix turn helix with hydrophobic interactions. First and second helices are involved in recognition of specific target genes, while the third helix makes direct contact with the major groove of its target DNA (Williams and Grotewold, 1997). Two cysteine residues in the third helix play a role in redox dependent DNA binding.

4.2.3 Identification and Sequence Analysis of OsDOF18

The protein sequence of ZmDOF1 (predicted by PLACE, Consite and JASPAR) when subjected to BLAST to identify similar DOF protein from O. sativa revealed its 54% sequence similarity with an uncharacterized protein. Protein functional analysis was performed using InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan5/) and this protein showed identity with InterPro accession number IPR003851 Znf_DOF, in this manner, confirming it as a DOF like proteins. This protein sequence was found in the Database of Rice Transcription Factors under locus ID Os08g38220 named as OsDOF18/MNB1A. The estimated molecular weight of OsDOF18 is 28.8 kDa and pI is 6.74.

Multiple alignment was carried out with DOF proteins from other plants that showed that it encodes a unique DNA binding domain of 55 amino acids containing a C2-C2 zinc finger (Figure 4). Shimufurutani et al. (1998) observed the involvement of
Figure 3: Protein sequence alignment of OsMYB1 with different MYB proteins and structure of DNA binding domain

MYB proteins from different plant members including MYB from A. thaliana, Z. mays, T. aestivum and V. vinifera were analyzed by TCOFFEE. Genbank accession numbers for OsMYB1, TaWRKY80, HvWRKY1, ShputativeWRKY, TaWRKY8, BdWRKY40 and ZmWRKY71 are EAY76504.1, BAA01730.1, NP_001141619.2, CBH32525.1, and XP_002276375.1 respectively. Conserved residues are shaded in different colors. Purple cylinders indicate the α helices of R2 and R3 domain. Asterisks indicate the conserved tryptophan residues important for folding of DNA binding domain. Arrow heads represents conserved cysteine residues involve in redox dependent DNA binding of MYB1 with DNA.
Figure 4: Protein sequence alignment of OsDOF18 with DOF proteins from different plants and structure of DNA binding domain

DOF proteins from different plant members including DOF from Z. mays, and S. bicolor were analyzed by TCOFFEE. Genbank accession numbers for OsDOF18, ShDOF22, ShDOF7, ZmDOF1, and ZmDOF2 are NP_001062110, DAA34025, AGL40014, ABF51012, and DAA40325 respectively. Conserved residues are shaded in different colors. Purple cylinders indicate the predicted α helices and green arrow represents predicted β strand in DNA binding domain. Asteriks indicate the conserved cysteine residues that make contact with Zn to form cys/cys Zinc finger motif. Middle cysteine presented by arowhead is involved in negative regulation of DNA binding. Aromatic residues (tyrosine, tryptophan) represented by arrowheads play important role in DNA binding (Shimufurutani et al., 1998).
middle cysteine residue in negative regulation of DNA binding and aromatic amino acids, tyrosine and tryptophan in binding with DNA.

4.3 CLONING, OVEREXPRESSION AND PURIFICATION OF OsWRKY71, OsDOF18 AND OsMYB1 DNA BINDING DOMAINS

To confirm the interaction of in silico identified target transcription factors with particular regulatory motifs in OsRGLP2 promoter, DNA binding domains of OsWRKY71, OsDOF18 and OsMYB1 were cloned, expressed and purified from E. coli. For this purpose, Glutathione S-transferase (GST) Gene Fusion System was used.

Proteins fused to GST have been reported to be successfully used in DNA-protein interaction studies (Vikis and Guan, 2004). From pGEX vector series, pGEX 4T-1 vector was used. The expression of pGEX4T-1 is under the control of the tac promoter (hybrid of trp and lac promoter), and expression is provoked using lactose analog isopropyl β-D-thiogalactoside (IPTG). These vectors also contain lacI° gene, the product of which is a repressor protein. The repressor protein interacts with the operator region of the tac promoter and prevent its expression until it is induced by IPTG, thus upholding firm regulation on expression of the sequence inserted into vector (Figure 5).

4.3.1 Cloning of DNA Binding Domains

Total RNA was extracted from 10 days old seedlings of KS282 (Oryza sativa L. SspIndica). Isolated total RNA was run on Agarose gel to check the quantity and quality of RNA (Plate 1) and later on, subjected to cDNA synthesis using oligo (dT) primers from RNA. cDNA was used as template to amplify the DNA binding domains of OsWRKY71 and OsMYB1 while OsDOF18 DNA binding domain was amplified from cDNA clone obtained from NIAS databank Japan.
Figure 5: Expression vector pGEX-4T1 for *OsWRKY71*, *OsMYB1* and *OsDOF18*

Plate 1: RNA isolated from 10 day old seedlings

RNA was isolated from 10 days old rice seedlings. **Lane M:** 1Kb DNA marker. **Lane 1-7:** RNA from rice seedlings. Two bands represents 18S and 28S rRNAs
4.3.1.1 Cloning of OsWRKY71 DNA binding domain

A 327 bp fragment of the OsWRKY71 DNA binding domain was amplified from rice cDNA (Plate 2A), purified from gel and ligated with BamHI and XhoI digested pGEX4t-1 vector. The ligation mixture was transformed into DH5α and grown overnight on LB plates. DH5α is highly transformable recA- strain, which means that it allows stable maintenance of plasmid DNA due to low recombination rate. Colony PCR was performed on selected 10 colonies and presence of 327 bp band confirmed the positive clones (Plate 2B). Plasmids extracted from overnight grown cultures of positive colonies were digested with BamHI and XhoI and checked on 1% agarose gel. The presence of 327 bp band demonstrated the successful ligation and transformation (Plate 2C). The upper ~4.5 Kb fragment in lane 1 was of linearized pGEX-4T 1.

4.3.1.2 Cloning of OsMYB1 DNA binding domain

It has been reported that shortening of MYB proteins outside the MYB domain does not affect binding specificity (Ramsay et al., 1992). Hence only OsMYB1 DNA binding domain with some flanking sequence was PCR amplified, and analyzed on agarose gel along with 1 Kb molecular weight marker. The PCR amplified product showed one clear band of 417 bp (Plate 3A). To clone this amplified DNA binding domain, gel band was purified and ligated into predigested pGEX-4T1 vector followed by transformation into DH5α cells. The clones were screened for the presence of insert by colony PCR and positive clones were selected. Colony PCR showed sharp and the single clear band as shown in Plate 3B. Plasmid was isolated from positive clones and subjected to restriction digestion by BamHI and XhoI. The restriction digestion products of pGEX-OsMYB1 are shown in Plate 3C. After confirmation of correct digest product size, insert sequence was further confirmed by commercial sequencing (Eurofins) of plasmid with pGEX-forward primer.
Plate 2: Cloning of *OsWRKY71* DNA binding domain

**A:** PCR optimization of *OsWRKY71* at 66°C. **B:** Colony PCR (Lane M: 1 Kb DNA ladder, Lane 1-10: Colonies 1-10, Lane 11: -ve control). **C:** Restriction digestion of pGEX-*OsWRKY71* (Lane 1: Restricted plasmid, Lane 2: Uncut plasmid)
Plate 3: Cloning of OsMYB1 DNA binding domain

A: PCR optimization of OsMYB1 at 62°C. B: Colony PCR (Lane M: 1 Kb DNA ladder, Lane 1-8: Colonies 1-8, Lane 9: -ve control). C: Restriction digestion of pGEX-OsMYB1 (Lane 1: Uncut plasmid, Lane 2: Restricted plasmid).
4.3.1.3 Cloning of OsDOF18 DNA binding domain

An attempt was made to construct an expression vector containing the full coding sequence of OsDOF18. It ended in failure, perhaps because the expression of the whole coding region of DOF protein is leaky and toxic for E. coli cells. Or it may have produced degraded forms of DOF18 proteins. However, a plasmid that allowed the expression of DOF18 DNA binding domain only with some flanking region was constructed. A 453 bp product of OsDOF18 DNA binding domain containing sequence was successfully amplified by PCR using gene specific primer sets (Plate 4A) and cloned into pGEX-4T1. The clones were confirmed by colony PCR for the presence of insert (Plate 4B). The plasmid was isolated from positive clones and subjected to restriction digestion by BamHI and XhoI. The restriction digestion products of pGEX-OsDOF18 are shown in Plate 4C. After confirmation of correct digest product size, insert sequence was further confirmed by commercial sequencing (Eurofins) of plasmid with pGEX-forward primer.

4.3.2 Expression Level of Recombinant GST-OsWRKY71 (167-271) in Different E. coli Strains

To investigate the expression level of GST-OsWRKY71 (167-271), pGEX-OsWRKY71 was transformed into three E. coli strains, BL21 (DE3), BL21 (DE3) - pLysS and BL21-CodonPlus (DE3)-RP and protein expression was induced. BL21 (DE3) FompT hsdSB (rB mB ) gal dcm (DE3) is Lon and ompT proteases deficient strain and is used for the expression of non-toxic genes. BL21 (DE3) pLysS cells (ompT hsdSB (rB mB ) gal dcm) (DE3) pLysS (Cm^R) has pLysS plasmid which produces T7 lysozyme and decreases the basal level expression of the fusion protein. This strain is used for the heterologous expression of those genes which are considered toxic for E. coli.
Plate 4: Cloning of OsDOF18 DNA binding domain

A: PCR optimization of OsDOF18 at 65°C. B: Colony PCR (Lane M: 1 Kb DNA ladder, Lane 1-7: Colonies 1-7, Lane 8: -ve control). C: Restriction digestion of pGEX-OsDOF18 (Lane 1: Uncut plasmid, Lane 2: Restricted plasmid)
BL21-CodonPlus (DE3)-RP strain (BF\textit{ompT hsdSB}_{rB\_mB} dcm + Tet gal1
(DE3) \textit{endA Hte [arg UproL Cam]}) strain is chosen in cases of codon bias. High-level
expression of a heterologous recombinant gene with codons that are rarely used by \textit{E. coli}
can generate depletion of the internal tRNA pool. The BL21-CodonPlus (DE3)-RP strain contains extra copies of uncommon \textit{E. coli} tRNA genes for the arginine codons AGA and AGG, plus the proline codon CCC. The same quantity of all three \textit{E. coli} strains cell lysates was examined by SDS-PAGE to check the protein expression pattern. After IPTG induction, all three \textit{E. coli} strains were capable to produce GST-WRKY71 (167-271) in 4 hours (Plate 5). BL21(DE3) showed more expression of GST-WRKY71 (167-271) than the protein produced by BL21-CodonPlus (DE3)-RP and BL21 (DE3)-pLysS (Plate 5A). BL21-CodonPlus (DE3)-RP and BL21 (DE3)-pLysS showed some contamination of GST which might be due to proteolysis of the translated products. These strains also showed considered dissimilarity in their growth profile. The OD\textsubscript{600} value achieved by BL21 (DE3) was higher as compared to BL21 (DE3)-pLysS and BL21-CodonPlus (DE3)-RP cells. This may be clarified in connection with “protein burden” in BL21 (DE3) pLysS and BL21-CodonPlus (DE3)-RP cells when GST-\textit{OsWRKY71} (167-271) may be produced at a comparatively elevated level in the beginning of IPTG induction and this protein load might ultimately force BL21(DE3)-pLysS and BL21-CodonPlus (DE3)-RP cells into growth arrest.

4.3.3 Purification of GST-\textit{OsWRKY71} from BL21

Bacterial culture was sonicated and centrifuged. The supernatant and pellet obtained after sonication were loaded on SDS PAGE to check the solubility of GST-\textit{OsWRKY71} (167-271). Results demonstrated that the solubility of the GST-\textit{OsWRKY
Plate 5: Expression and purification of GST-OsWRKY71 (167-271)

A): Expression of GST-OsWRKY71(167-271) in BL21, BL21 pLysS, and BL21 Codon Plus. **Lane 1:** Uninduced BL21, **Lane 2:** Induced BL21, **Lane 3:** Uninduced BL21 pLysS, **Lane 4:** Induced BL21- pLysS, **Lane 5:** Uninduced BL21-C+, **Lane 6:** Induced BL21-C+. B): Purification of expressed GST-OsWRKY71 (167-271) from BL21. **Lane 1:** Uninduced BL21, **Lane 2:** Induced BL21, **Lane 3:** Sonicated supernatant, **Lane 4:** Sonicated Pellet, **Lane 5:** Flowthrough, **Lane 6-7:** Wash 1, 2, **Lane 8-14:** Elutions 1-7. C): Purification of GST-OsWRKY71 (167-271) by cation exchange chromatography. **Lane 1-8:** Fractions containing purified GST-OsWRKY71 (167-271).
OsWRKY71 (167-271) was expressed with N-terminal GST tag and was purified using GST affinity columns. After protein purification by affinity chromatography, the eluted soluble GST-OsWRKY71 (167-271) protein was checked by measuring the absorbance at OD$_{280}$ and then analyzed by running on SDS-PAGE. A classic elution form of the protein fractions collected from a GST affinity column is shown in Plate 5B. Fractions 2-7 contained the eluted GST-OsWRKY71 (167-271) protein from the GST column. The specific 38 kDa band eluted in fractions was approximately 70% purified and contained some nonspecific bands.

Elutions containing GST-OsWRKY71 (167-271) were subjected to cation exchange chromatography for further purification on the basis of charge (Plate 5C). For this purpose, GST-OsWRKY71 (167-271) was dialyzed in phosphate buffer (pH 7) as protein has a pI of 8.25. At a pH below the pI of protein, the net charge of GST-OsWRKY71(167-271) was positive and it bound to the cation exchanger reversibly by electrostatic forces. Then, GST-OsWRKY71 (167-271) was eluted by disturbing the electrostatic force between column and protein by providing NaCl gradient from 100 mM to 1M.

4.3.4 Expression Level of Recombinant GST-OsDOF18 (1-148) in Different E. coli Strains

Expression of pGEX-OsDOF18(1-148) was carried out in three E. coli strains and protein expression level was compared (Plate 6A). The expression pattern of GST-OsDOF18 (1-148) in all three strains is shown in Plate 6. With BL21 (DE3) cells, a significant amount of GST-OsDOF18 (1-148) was observed in cell lysate after IPTG induction for 4 hours. When BL21 (DE3)-pLysS cells harboring pGEX-OsDOF18
Plate 6: Expression and purification of GST-OsDOF18 (1-148)

A: Expression of GST-OsDOF18 (1-148) in BL21, BL21 pLysS, and BL21 Codon Plus. **Lane 1:** Uninduced BL21, **Lane 2:** Induced BL21, **Lane 3:** Uninduced BL21 pLysS, **Lane 4:** Induced BL21-pLysS, **Lane 5:** Uninduced BL21-C+, **Lane 6:** Induced BL21-C+.  

**B:** Purification of Expressed GST-OsDOF18 (1-148) from BL21. **Lane 1:** Uninduced BL21, **Lane 2:** Induced BL21, **Lane 3:** Sonicated sample, **Lane 4:** Sonicated supernatant, **Lane 5:** Sonicated pellet, **Lane 6:** Flowthrough, **Lane 7:** Wash 1, **Lane 8-14:** Elutions: 1-7.  

**C:** Concentrated GST-OsDOF18.
were examined, expression pattern was poor. Protein products were not detectable on SDS PAGE, no matter whether or not the GST-DOF18 (1-148) was being expressed. In addition, when BL21-CodonPlus(DE3)-RP were transformed with pGEX-OsDOF18, 42kDa specific protein band was detectable on SDS PAGE. Growth profile was higher in BL21 (DE3) cells as compared to the other two strains. BL21 (DE3) strain was, therefore, chosen for large scale expression and purification.

4.3.5 Purification of GST-OsDOF18 (1-148)

Overexpressed GST-OsDOF18 (1-148) was purified by using glutathione Sepharose affinity purification method. GST-OsDOF18 (1-148) bound to glutathione Sepharose column was eluted with 50 mM reduced glutathione. Elutions were clear, but contained some nonspecific proteins from E. coli cells. GST-OsDOF18 (1-148) was further purified by cation exchange chromatography. Fractions of purified protein were analyzed on SDS PAGE (Plate 6C), pooled and concentrated for EMSA.

4.3.6 Expression Level of Recombinant GST-OsMYB1(88-222) Protein in Different E. coli Strains

pGEX-OsMYB1 was transformed in all three E. coli strains to examine which of the three strains was able to demonstrate better expression. It was found that expression in BL21 (DE3) cells was significantly higher than BL21-CodonPlus (DE3)-pLysS and no overexpression of recombinant protein was detected with BL21 (DE3)-pLysS. BL21 (DE3) has also showed a higher growth rate in comparison to BL21(DE3)-pLysS and BL21-CodonPlus (DE3)-RP. It was deduced that BL21 (DE3) is the ideal option for expressing pGEX-OsMYB1 (88-222) (Plate 7A).

4.3.7 Purification of GST-OsMYB1

Overexpressed GST-OsMYB1 (88-222) was purified using the glutathione
Plate 7: Expression and purification of GST-OsMYB1 (88-222).

Sepharose affinity purification method. GST-OsMYB1 (88-222) bound to glutathione Sepharose column was eluted with 50 mM reduced glutathione (Plate 7B). Elutions were cleared, but contained some nonspecific proteins from *E. coli* cells. GST-OsMYB1 (88-222) was further purified by cation exchange chromatography and purified protein was checked on SDS PAGE (Plate 7C). Protein was 95% pure now, fractions were pooled and concentrated by using amplicon 3K tubes and then used for downstream characterization.

4.4 MALDI-TOF MASS SPECTROMETRY (MS) OF CLEAVED PROTEINS FRAGMENTS

In order to confirm that purified proteins are specific ones, GST tag was removed from *Os*WRKY71, *Os*DOF18 and *Os*MYB1 DNA binding domains by on column cleavage with thrombin. Then cleaved peptides were subjected to MALDI-TOF analysis. *Os*WRKY71 has one cleavage site for thrombin and it was predicted to be cleaved into two fragments of 26000 Da and 11961.82 Da representing GST and *Os*WRKY71 (167-271) respectively. The cleaved *Os*WRKY71 fragment was purified from cleaved GST and then subjected to MS analysis. Mass spectra for *Os*WRKY71 contained a fragment of 11961.83 Da size verifying the presence of specific *Os*WRKY71 peptide (Figure 6).

For *Os*MYB1, there were two cleavage sites for thrombin which can result in generating three fragments; one of 26000 Da for GST, and two fragments of 4986.56 Da and 10201.74 Da for *Os*MYB1 peptides (Figure 7).

The size of GST-**Os**DOF18 was about 42 KDa. Thrombin cleavage resulted in cleaving the GST-DOF18 into two fragments, 26000 Da for GST and 15644.37 Da. The Mass spectrometer used was not able to be calibrated for large size peptides and proteins, so no specific spectra were observed for *Os*DOF18 peptide. There were
Figure 6: MALDI-TOF spectrum of the cleaved OsWRKY71 (167-271)

Signals that matched the theoretical peptide masses of OsWRKY71 fragment are labeled with m/z values of 11961.83.
Figure 7: MALDI-TOF spectra of the cleaved OsMYB1 fragments

A: Represents the spectra of 48 a.a fragment of OsMYB1. B: Mass spectra of the 90 amino acids residues of OsMYB1. Signals that matched the approximate theoretical peptide masses of OsWRKY71 fragment are labeled with m/z values.
spectra upto 12000 Da peptides range.

4.5 BINDING OF RECOMBINANT PROTEINS WITH OsRGLP2 PROMOTER REGION

To test the hypothesis of the regulation of the OsRGLP2 promoter by OsWRKY71, OsMYB1 and OsDOF18 proteins, Electrophoretic Mobility Shift Assays (EMSAs) were carried out to assess the potential binding of these proteins to OsRGLP2 promoter region.

4.5.1 Labelling Efficiency

Non-radioactive EMSA has many benefits over classical radioactive method. Digoxigenin labeled probes are more secure to use in the laboratory and these are more stable as compared to radioactive labeled probes. This technique is also competitive in sensitivity, time and cost. Efficiency of labeled reactions was assessed by generating a dot blot. Serial dilution was applied on positively charged nylon membrane and detection was carried out. The resulting intensity was then compared to the control labeled oligonucleotides provided with the kit (Plate 8).

4.5.2 Analysis of Interaction of The GST-OsWRKY71 (167-271) with Putative WRKY Binding Motifs on OsRGLP2 Promoter

WRKY proteins are generally considered to bind with the consensus true W-box sequence (C/T)TGAC(C/T) (Eulgem and Somssich, 2007). The W-box is present in the promoter regions of a lot of plant defense providing genes (Eulgem et al., 2000), including WRKY itself. Tobacco CHN50 promoter contains various W-boxes, and different WRKY proteins interact with these boxes (Yang et al., 1999). A current microarray analysis discovered that the promoters of defense related genes contain W-box as an important element, for instance, occurrence of W-boxes in the promoters of PR-1, a marker gene for SAR (Maleck et al., 2000).
Plate 8: Labelling efficiency of DIG labeled oligos using a dot blot

Different dilutions of labeled *OsWRKY71*-1, *OsDOF18*-1 and *OsMYB1*-1 oligos were spotted on membrane and compared with control reaction. 50 fmol was chosen for EMSA experiments.
There are four putative W-boxes on OsRGLP2 promoter: two TTGACT W-boxes on positive strand and two CTGACC W-Boxes on the negative strand. Although numerous TGAC core sequences are also found, WRKY proteins have also been found to interact with convergent repeats of TGAC (Wang et al., 1998). Both computational and functional studies have found clusters of W-boxes in the promoters of stress-inducible genes (Maleck et al., 2000). Multiple W-boxes have a synergistic effect on transcription as observed by Mare et al. (2004).

To characterize the OsWRKY71 binding sites in OsRGLP2 promoter, three double stranded 30 bp promoter fragments containing either true W-box or core W-box sequences were prepared as probe and examined for their affinity with recombinant GST-OsWRKY71 (167-271).

4.5.2.1 Probe containing one true and one core W-Box

After DIG labeling, the fragments were examined for their capability to interact with purified GST-OsWRKY71 (167-271) using EMSAs. Competition assays were performed with wild-type and mutant DNA probes. OsWRKY71 probe 1 from nucleotide -635 to -606, contains TTGACT true W-box sequence and one TGAC core sequence. Purified GST protein was used as negative control. No binding was observed for two negative controls: reaction containing no protein and reaction containing GST protein only with labeled oligos, (Plate 9A). GST-OsWRKY71 (167-271) bound to the wild-type WRKY71-probe1 and to verify the binding specificity, competition experiment was done with 200 fold excess of unlabeled fragment. Evidently, addition of 200-fold excess unlabeled fragment entirely outcompeted the interaction of GST-OsWRKY71 with the probe, suggesting that this OsRGLP2
promoter fragment specifically bound to WRKY71. No shift was observed when true W-box and core TGAC sequences were mutated to AAAA. When the competitor analysis was performed by mutating both binding sites in the WRKY71-1 probe, it was not able to eliminate the interaction of the protein with the probe.

The interaction of rice nuclear proteins with the W-box was also investigated by EMSAs. The shifts were observed as multiple bands on the EMSA gels (Plate 9B), which is most likely the result of binding of different W-box-binding proteins to the W-box element. Multiple shifts might result due to binding of other binding proteins to the OsWRKY71-1 oligos as it also contains binding sites for other proteins. On the other hand, as WRKY proteins function by interacting with other WRKY proteins and other protein partners, the several bands observed might imply the regulation via a complex transcriptional mechanism.

4.5.2.2 Probe containing one true W-Box

WRKY71-probe 2 ranging from nucleotide -781 to -752, contains one true W-Box sequence. No shift was observed in lanes containing labeled probe without protein and labeled probe with GST. GST-OsWRKY71 (167-271) bound directly to WRKY71-Probe2 when incubated with labeled probe (Plate 10A). To confirm binding, W-box sequence was mutated to AAAA and performed EMSA by using labeled mutated probe. EMSA probe that contained mutated W-box failed to bind with GST-OsWRKY71. There was no shift in the negative control of labeled mutant probe. As competition, 200 fold molar excess of unlabeled cold WRKY71-probe 2 was added to the EMSA reaction before the addition of labeled probe. Intriguingly cold EMSA unlabeled probe competed well with radiolabeled probe. When an excess of the mutated cold EMSA probe was added to the reaction, it failed to compete with the
Plate 9: In vitro DNA binding ability of *OsWRKY71* with *OsRGLP2* fragment containing one core W-box and one true W box

A: EMSA to test the binding ability of recombinant *OsWRKY71* with DIG labeled *OsRGLP2* fragment. **Lane 1:** *OsWRKY71* -1 oligos negative control, **Lane 2:** GST negative control, **Lane 3:** 1 µg GST-*OsWRKY71* with *OsWRKY71* -1 oligos, **Lane 4:** with 200 molar excess of unlabeled *OsWRKY71* -1 oligos, **Lane 5:** With 200 fold molar excess of unlabeled mutant oligos, **Lane 6:** 500 fold molar excess of unlabeled mutant oligos, **Lane 7:** With mutant labeled oligos, **Lane 8:** Mutant labeled oligos negative control. B: EMSA reaction of *OsRGLP2* fragment with nuclear protein extract. **Lane 1:** 7 µg nuclear proteins extract with labeled *OsWRKY71* -1 oligos, **Lane 2:** With 200 molar excess of unlabeled *OsWRKY71* -1 oligos.
Plate 10: *In vitro* DNA binding ability of *OsWRKY71* with *OsRGLP2* fragment containing one true W box

A: EMSA to test the binding ability of recombinant *OsWRKY71* with DIG labeled *OsRGLP2* fragment. **Lane 1:** *OsWRKY71*-2 oligos negative control, **Lane 2:** GST negative control, **Lane 3:** 1 µg GST-*OsWRKY71* with *OsWRKY71*-2 oligos, **Lane 4:** with 200 molar excess of unlabeled *OsWRKY71*-2 oligos, **Lane 5:** With mutant labeled oligos, **Lane 6:** 200 fold molar excess of unlabeled mutant oligos, **Lane 7:** Mutant labeled oligos negative control

B: EMSA reaction of *OsRGLP2* fragment with nuclear protein extract. **Lane 1:** 7 µg nuclear proteins extract with labeled *OsWRKY71*-2 oligos, **Lane 2:** With 200 molar excess of unlabeled *OsWRKY71*-2 oligos.
labeled W-box containing fragment of OsRGLP2 promoter.

When the interaction of rice nuclear proteins with the WRKY-probe 2 containing CTGACC was investigated by EMSA, multiple shifts were observed as observed in the case of OsWRKY71-1 probe (Plate 10B).

**4.5.2.3 Probe containing one core W-Box TGAC on both strands**

OsRGLP2 promoter contains a number of putative core W-box sequence TGAC. To determine whether purified GST-OsWRKY71 (167-271) bind with core sequence, an OsRGLP2 fragment was chosen as probe which contained TGACT sequence. There was no shift in both negative controls i.e labeled probe without protein and labeled probe with GST. GST-OsWRKY71 (167-271) bound to core sequence and a shift was observed, but this binding was weak as compared to true W-box sequence (Plate 11A). The specificity of the shift was further established through competition and mutation assays. Shifted band was totally competed with 200 fold molar excess of unlabeled cold probe. It was observed that there was no shift when the core sequence was mutated to AAAA. Excess of the unlabeled mutant probe also failed to abolish the WRKY-DNA complex with labeled OsWRKY71-Probe 3.

In parallel, DNA binding capabilities of rice nuclear proteins with the W-box was investigated by EMSAs. Results have shown that migration of WRKY-Probe 3 is strongly retarded in the presence of the nuclear protein extract. More than one shift was observed (Plate 11B) which is most likely due to the interaction of different WRKY proteins or other proteins with the W-box element. In competition assay, 200 fold molar excess of unlabeled specific competitor diminished the shifted signals suggesting the specific interactions.
Plate 11: *In vitro* DNA binding ability of OsWRKY71 with OsRGLP2 fragment containing one core W-box on both strands

**A:** EMSA to test the binding ability of recombinant OsWRKY71 with DIG labeled OsRGLP2 fragment. **Lane 1:** OsWRKY71-3 oligos negative control, **Lane 2:** GST negative control, **Lane 3:** 1 µg GST-OsWRKY71 with OsWRKY71-3 oligos, **Lane 4:** with 200 molar excess of unlabeled OsWRKY71-3 oligos, **Lane 5:** With mutant labeled oligos, **Lane 6:** 200 fold molar excess of unlabeled mutant oligos, **Lane 7:** Mutant labeled oligos negative control. **B:** EMSA reaction of OsRGLP2 fragment with nuclear protein extract. **Lane 1:** 7 µg nuclear proteins extract with labeled OsWRKY71-3 oligos, **Lane 2:** With 200 molar excess of unlabeled OsWRKY71-3 oligos.
As a whole, this data showed that OsWRKY71 can identify and bind with the OsRGLP2 promoter \textit{in vitro}, interacting specifically with core W-box (TGAC) containing cis-regulatory elements. Present study demonstrated that a W-box core sequence (TGAC), but not an extended W-box (C/TTGACC/T), is the interacting site of the trans-activating factor(s) that control(s) the expression of OsRGLP2 gene but adjacent sequence increased the binding affinity.

4.5.3 Analysis of Interaction of the GST-MYB1 (88-222) With Putative MYB Binding Motifs on OsRGLP2 Promoter

It has been accounted that MYB transcription factors can recognize a specific binding motif, A/TAACCA and C/TAACG/TG (Abe \textit{et al.}, 2003). In OsRGLP2 promoter, MYB-binding motifs (5’-TAACCA- 3’ and 5’-AAACCA- 3’) were found. The objective of these experiments was to examine whether the OsMYB1 protein binds to the different putative MYB responsive motifs identified in the OsRGLP2 promoter.

4.5.3.1 Probe containing TAACCA motif

The \textit{cis}-regulatory elements for binding of MYB proteins in the promoter regions of many genes all containing the same core sequence of AACC but differ in binding specificity due to differences in the adjacent nucleotides of MYB motif. MYB2 protein showed interaction with CTAACCA consensus sequence containing AACC core sequence of the MYB1 ATcis-element (Abe \textit{et al.}, 1997). A 30 bp fragment of OsRGLP2 promoter containing TAACCA motif was DIG labeled and incubated with purified GST-OsMYB1. Recombinant GST-OsMYB1 bound to the TAACCA motif in the OsRGLP2 promoter, but GST alone did not bind (Plate 12A). Protein-DNA complex was reduced by the addition of 200 fold molar excess of unlabeled
competitor oligos. The DNA binding activity of *OsMYB1* abolished when protein was incubated with mutant labeled oligos. Two hundred fold excess of unlabeled mutant oligos did not affect the binding of GST-*OsMYB1* with TAACCA sequence.

Analysis of nuclear protein interaction with MYB1-Probe 1 by EMSA demonstrated that nuclear proteins were able to bind with *OsRGLP2* promoter fragment. This binding was reduced by the addition of 200 fold molar excess of unlabeled competitor oligos suggesting that interaction was specific (Plate 12B).

**4.5.3.2 Probe containing AAACCA motif**

As shown in Plate 13A, the GST-*OsMYB1* fusion protein was able to bind to the DIG-labeled MYB1-Probe 2 containing AAACCA motif, predicted binding site for MYB protein. The DIG-labeled probe competed with unlabeled cold competitor probes, demonstrating that the signal from the probe represents the specific binding ability of the *OsMYB1* with the *OsRGLP2* promoter fragment. When the MYB box was mutated to TCATGA, no shift was observed. Additionally, probes with a mutation in the MYB core sequence was unable to compete with the wild-type probe. These results demonstrated that *OsMYB1* was able to directly and specifically bind to the MYB-binding site of the *OsRGLP2* promoter *in vitro*. These results suggest that *OsMYB1* can bind specifically to *OsRGLP2* promoter A/TAACCA sequences *in vitro*.

Analysis of nuclear protein interaction with MYB1-Probe 2 by EMSA demonstrated that nuclear proteins were able to bind with *OsRGLP2* promoter fragment. This binding was reduced by the addition of 200 fold molar excess of unlabeled competitor oligos, suggesting that interaction of nuclear protein (s) with probe was specific (Plate 13B).
Plate 12: *In vitro* DNA binding ability of *OsMYB1* with *OsRGLP2* fragment containing one MYB binding site TAACCA

A: EMSA to test the binding ability of recombinant *OsMYB1* with DIG labeled *OsRGLP2* fragment. **Lane 1:** *OsMYB1*-2 oligos negative control, **Lane 2:** GST negative control, **Lane 3:** 1 µg GST-*OsMYB1* with *OsMYB1*-2 oligos, **Lane 4:** with 200 molar excess of unlabeled *OsMYB1*-2 oligos, **Lane 5:** With mutant labeled oligos, **Lane 6:** 200 fold molar excess of unlabeled mutant oligos, **Lane 7:** Mutant labeled oligos negative control  

B: EMSA reaction of *OsRGLP2* fragment with nuclear protein extract. **Lane 1:** 7 µg nuclear proteins extract with labeled *OsMYB1*-2 oligos, **Lane 2:** With 200 molar excess of unlabeled *OsMYB1*-2 oligos.
**Plate 13:** *In vitro* DNA binding ability of *OsMYB1* with *OsRGLP2* fragment containing one MYB binding site AAACCA

**A:** EMSA to test the binding ability of recombinant *OsMYB1* with DIG labeled *OsRGLP2* fragment. **Lane 1:** *OsMYB1*-1 oligos negative control, **Lane 2:** GST negative control, **Lane 3:** 1 µg GST-*OsMYB1* with *OsMYB1*-1 oligos, **Lane 4:** with 200 molar excess of unlabeled *OsMYB1*-1 oligos, **Lane 5:** With mutant labeled oligos, **Lane 6:** 200 fold molar excess of unlabeled mutant oligos, **Lane 7:** Mutant labeled oligos negative control

**B:** EMSA reaction of *OsRGLP2* fragment with nuclear protein extract. **Lane 1:** 7 µg nuclear proteins extract with labeled *OsMYB1*-1 oligos, **Lane 2:** With 200 molar excess of unlabeled *OsMYB1*-1 oligos.
4.5.4 Analysis of the Interaction of The GST-DOF (1-148) With Putative DOF Binding Motifs on OsRGLP2 Promoter

To investigate the function of the several AAAG motifs in the OsRGLP2 5’-flanking region, EMSAs were performed using double-stranded DIG-labeled DNA Probes 1 and 2 representing two and one DOF boxes respectively.

4.5.4.1 Probe containing two AAAG motifs

To analyze if GST-OsDOF18 is able to specifically interact with 30 bp fragment of OsRGLP2 promoter containing 2 AAAG motifs, EMSA was performed. GST-DOF18 efficiently bound to the probe when incubated with DIG labeled DOF-Probe1 and this complex went away when protein was incubated with 200 fold molar excess of same unlabeled oligo. To investigate whether DOF protein binding to AAAG motif is specific or not, a mutant version of the probe was created in which AAAG sequence was mutated to AGAC. EMSA with mutant probe did not result in any shift (Plate 14A). As predicted, this shift was not observed when the GST protein was used in this experiment.

The interactions of rice nuclear proteins with the W-box was also investigated by EMSAs. The shifts were observed as multiple bands on the EMSA gels (Plate 14B), which is most likely the result of binding of different binding proteins to the AAAG element.

4.5.4.2 Probe containing one AAAG motif

When DOF-Probe 2 containing one AAAG motif was incubated with GST-OsDOF18 protein, a shift was observed. The binding specificity was established using
Plate 14: *In vitro* DNA binding ability of *OsDOF18* with *OsRGLP2* fragment containing two AAAG boxes

**A:** EMSA to test the binding ability of recombinant *OsDOF18* with DIG labeled *OsRGLP2* fragment. **Lane 1:** *OsDOF18*-1 oligos negative control, **Lane 2:** GST negative control, **Lane 3:** 1 µg GST-*OsDOF18* with *OsDOF18*-1 oligos, **Lane 4:** with 200 molar excess of unlabelled *OsDOF18*-1 oligos, **Lane 5:** With mutant labeled oligos, **Lane 6:** 200 fold molar excess of unlabelled mutant oligos, **Lane 7:** Mutant labeled oligos negative control. **B:** EMSA reaction of *OsRGLP2* fragment with nuclear protein extract. **Lane 1:** 7 µg nuclear proteins extract with labeled *OsDOF18*-1 oligos, **Lane 2:** With 200 molar excess of unlabelled *OsDOF18*-1 oligos.
modified probes. The interaction of the GST-OsDOF18 to the DOF-box in the OsRGLP2 gene promoter was eliminated when the core AAAG box was mutated to AGAC (Mena et al., 1998). As shown in Plate 15A, mutant version of the DOF probe was not able to interact with GST-OsDOF18 protein or not capable to abolish the shift corresponding to the interaction of GST-OsDOF18 with wild-type probes, even at 200 molar excess.

These results suggest that DOF18 can bind to both an AAAG motif and a complex of two AAAG motifs. Since multiple DOF18 binding sites have been identified in OsRGLP2 promoter, DOF18 might bind to target promoter at multiple sites to exert its effect.

The interactions of rice nuclear proteins with the AAAG motif were also investigated by EMSAs. The shifts were observed as multiple bands on the EMSA gels (Plate 15B), which is most likely the result of binding of different binding proteins to the AAAG element. When 200 fold molar excess of unlabeled OsDOF18-2 oligos were added, one upper band disappeared while the other band was still present. This demonstrated the specificity of the upper band while the second band was nonspecific.

4.6 EXPRESSION ANALYSIS OF OsWRKY71, OsMYB1 AND OsDOF18 BY REAL TIME PCR

4.6.1 OsWRKY71 Gene Expression in Response to Various Abiotic Stresses

To investigate the stress responses of OsWRKY71 transcripts, an examination of global expression was performed using the Genevestigator. To evaluate the OsWRKY71 response to different stresses, gene expression was checked in rice seedlings using salt stress, drought, cold, high temperature and wounding. The relative expression of OsWRKY71 in reaction to a variety of abiotic stresses is shown in the
Plate 15: *In vitro* DNA binding ability of *OsDOF18* with *OsRGLP2* fragment containing one AAAG motif

**A**: EMSA to test the binding ability of recombinant *OsDOF18* with DIG labeled *OsRGLP2* fragment. **Lane 1**: *OsDOF18*-2 oligos negative control, **Lane 2**: GST negative control, **Lane 3**: 1 µg GST-*OsDOF18* with *OsDOF18*-2 oligos, **Lane 4**: with 200 molar excess of unlabeled *OsDOF18*-2 oligos, **Lane 5**: With mutant labeled oligos, **Lane 6**: 200 fold molar excess of unlabeled mutant oligos, **Lane 7**: Mutant labeled oligos negative control **B**: EMSA reaction of *OsRGLP2* fragment with nuclear protein extract. **Lane 1**: 7 µg nuclear proteins extract with labeled *OsDOF18*-2 oligos, **Lane 2**: With 200 molar excess of unlabeled *OsDOF18*-2 oligos.
The transcript level of OsWRKY71 gene was significantly up-regulated by cold, salt, drought and wounding stresses. OsWRKY71 expression was slightly induced (2 fold) by wounding and drought. In particular, OsWRKY71 expression was greatly induced (8 fold) by salt stress (500 mM). OsWRKY71 transcript was induced up to 5 fold when cold stress was imposed. Heat stress resulted in no induction in OsWRKY71 transcript. These results strongly suggest that OsWRKY71 gene plays important roles in providing tolerance to various abiotic stresses.

The FaWRKY1 from strawberry showed accumulation in response to elicitors and wounding (Encinas-Villarejo et al., 2009). The JcWRKY expression was up-regulated by dehydration and salt stresses (Agarwal et al., 2014). Marchive et al. (2007) reported the buildup of VvWRKY1 transcript in reaction to hormones, wounding and hydrogen peroxide. The WRKY38 showed accumulation in response to dehydration and low temperature (Mare et al., 2004). AtWRKY25 and AtWRKY53 transcript level was enhanced in response to heat and salt stresses (Li et al., 2009; Li et al., 2011; Jiang et al., 2009). TcWRKY53 expression level was also increased by cold and salt treatments (Wei, 2008). This shows that WRKY genes are involved to be expressed under different abiotic stresses.

4.6.2 OsDOF18 Gene Expression in Response to Various Abiotic Stresses

The transcript level of OsDOF18 increased under different abiotic stress conditions. However, its relative expression varied depending upon the kind of stress. OsDOF18 gene expression was significantly induced by salt, drought and wounding stresses (Figure 9). Expression in salt stress was increased up to 4.2 fold in comparison to control.
Figure 8: Expression analysis of *OsWRKY71*

**A:** Microarray data of *OsWRKY71* expression in abiotic stresses from Genevestigator. **B:** Relative expression of *OsWRKY71* in abiotic stresses by Real-time PCR. Bars represent standard errors of the mean based on three independent experiments.
Figure 9: Expression analysis of OsDOF18

A: Microarray data of OsDOF18 expression in abiotic stresses from Genevestigator. B: Relative expression of OsDOF18 in abiotic stresses by Real-time PCR. Bars represent standard errors of the mean based on three independent experiments.
Drought stress resulted in 3.8 fold increase in OsDOF18 transcript level, whereas in wounding stress, OsDOF18 transcript was slightly increased. These results suggested that OsDOF18 gene expression was high in salt and drought stresses and there was no significant increase in heat and cold stresses.

Corrales et al. (2014) showed that tomato cycling DOF factor transcript was up-regulated in response to different abiotic stresses including salt, drought and threshold temperature suggesting its role in stress conditions. Ma et al. (2015) also observed the up-regulation of DOF genes against cold, salt, heat and drought stresses.

4.6.3 OsMYB1 Gene Expression in Response to Various Abiotic Stresses

To elucidate the expression patterns of OsMYB1 gene under abiotic stresses for instance, drought, cold, salinity, dehydration and wounding, qRT-PCR analyses were performed. The experimental approach for qRT-PCR analysis was a comparative expression profiling and analysis of relative fold increase in transcript level of the OsMYB1 gene under various abiotic stress conditions. It was observed that expression patterns of the OsMYB1 change in different abiotic stresses.

The transcript level of OsMYB1 gene increased under different environmental stress conditions, however, their relative expression varies depending on the kind of the stress. OsMYB1 was significantly increased by wounding, salt, drought and heat stresses (Figure 10). Upon wounding stress conditions, the expression of OsMYB1 was increased up to 23 fold induction compared to control. Mahmood (2007) and Mahmood et al. (2013) reported the wound activity of this promoter for the first time by checking GUS analysis and Hussain (2015) studied the transcriptional level of GUS gene under the control of OsRGLP2 promoter and deleted fragments, which indicated
Figure 10: Expression analysis of OsMYB1

A: Microarray data of OsMYB1 expression in abiotic stresses from Genevestigator. B: Relative expression of OsMYB1 in abiotic stresses by Real-time PCR. Bars represent standard errors of the mean based on three independent experiments.
a significant increase in expression level due to strong activity of full length and 5' deleted promoter.

Drought and heat stresses resulted in ~3 fold increase in OsMYB1 gene expression compared to control. Under salt stress, the expression of OsMYB1 was peaked up to ~3 fold. These results strongly suggest that the OsMYB1 gene may play important roles in controlling many abiotic stresses faced by plants, including wounding, salt and drought stresses. Different MYB proteins have been reported to participate in different stress response processes (Cheng et al., 2013). JcMYB1 showed an increase in transcript level by PEG, NaCl and cold stresses as well as by hormones (Liang et al., 2014). Katiyar et al. (2012) analyze the up-regulation of MYB genes in rice and found that 14 OsMYB genes showed an enhanced transcript level under abiotic stresses including drought, salt and cold stresses. Out of these 14 genes, transcript level of ten genes was significantly high in response to drought stress.

4.7 EXPRESSION PATTERNS OF GST-OsWRKY71, GST-OsDOF18 AND GST-OsMYB1 in E. COLI UNDER DIFFERENT ABIOTIC STRESSES

To observe the consequences of the overexpression, the selected transcription factors were transformed into E. coli under various environmental stresses, cultures of BL/pGEX4T-1, BL/OsWRKY71, BL/OsMYB1, and BL/OsDOF18 were diluted and applied on different LB plates.

In recent times, a number of plant genes and transcription factors have been evaluated for their stress analysis using heterologous expression system of E. coli. SbsI-1, a unique salt responsive gene, from S. brachiata displayed drought and salinity tolerance in E. coli cells as investigated by Yadav et al.(2012). Guo et al. (2010) observed the drought tolerance in E.coli cells transformed with the ThPOD3 from
*Tamarix hispida*. Reddy *et al.*, (2010) investigated the expression of a cytoplasmic Hsp70 in *E. coli* cells and observed the defensive chaperone action against damage brought about by heat and salt stress. LEA proteins from soybean improved tolerance to *E. coli* cells against salt stress (Lan *et al.*, 2005).

4.7.1 **Expression Pattern of GST-OsWRKY71 in *E. coli***

In order to assess the function of expressed *OsWRKY71* protein in salt stress condition, the effect of high NaCl concentration was examined. As shown in Plate 16, BL/pGEX4T-1, BL/OsWRKY71 containing cells have the same growth pattern on standard LB plates. When plates were supplemented with different concentrations of NaCl, the number of colonies was different on each plate. It was observed that the survival of BL/OsWRKY71 cells was considerably superior than BL/pGEX4T-1 cells under different concentrations of NaCl. When NaCl concentration was increased to 600 mM in the medium, there were small and few BL/OsWRKY71 colonies in comparison to control BL/pGEX4T-1 cells containing no colony.

To analyze the effect of *OsWRKY71* in *E. coli* under desiccation, different dilutions were spread on LB IPTG plates containing different concentrations of mannitol. The quantity of BL/OsWRKY71 was elevated compared to BL/pGEX4T-1 when LB plates were supplemented with 1M mannitol. This outcome suggested that the expression of the *OsWRKY71* gene expression has provided tolerance to *E. coli* cells against salt and drought stress.

With the aim of identifying the outcome of *OsWRKY71* overexpression on the growth of *E. coli* cells against low and high-temperature, cultures induced with IPTG were moved to -80°C and 50°C respectively. After employing the temperature stresses
Plate 16: Spot assay of BL/pGEX4T-1 and BL/OsWRKY71 recombinants

Transformed *E. coli* cells were subjected to different abiotic stresses. Induced cultures OD was adjusted to OD$_{600}$=1. Then 10 μL of 50-, 100- and 200-fold diluted bacterial suspension was spotted on LB plates containing 400, 500 and 600 mM NaCl for salt stress; 500, 800 mM and 1M mannitol for dessication. Samples were spotted after 2, 4, 6 and 8 hours of cold stress and after 1, 2 and 3 hours of heat stress.
for different time intervals, the number of cells were compared in diluted cultures. Number of control cells was less as compared to BL/OsWRKY71 but the growth rate was stagnant for both control and BL/OsWRKY71 cells after 2, 4, 6 and 8 hours of cold treatment. Even though the cell number was decreased in both cultures after heat stress, BL/OsWRKY71 cells were more in comparison to control cells. Most of the BL/pGEX4T-1 cells did not survive after 3 hours at 50°C, while less BL/OsWRKY71 cells died after 3 hours. These results suggested that the OsWRKY71 gene significantly increased the tolerance to low and high-temperature stresses.

4.7.2 Expression Pattern of GST-OsDOF18 in E. coli

pGEX4t-OsDOF1 was overexpressed in E. coli cells to assess its effect under different abiotic stresses. BL/OsDOF1 cells showed better tolerance in high salt and desiccation treatment as compared to vector alone. At low and high temperature, bacterial growth was similar in BL/OsDOF1 and BL/pGEX4T-1 cells (Plate 17). These results revealed that OsDOF1 gene significantly induce tolerance under salt and dehydration stresses in E. coli.

4.7.3 Expression Pattern of GST-OsMYB1 in E. coli

The OsMYB1 recombinant protein expressed in BL21 leads to a better growth under various salt concentrations and there were fewer cells at 600 mM NaCl as compared to control having no growth. At high temperature, rate of survival of BL/OsMYB1 was significantly more in comparison to control; however, similar growth was observed in desiccation and cold treatment compared to control (Plate 18). While OsMYB1 gene has induced tolerance in the presence of high salt and under high temperature in E.coli.
Plate 17: Spot assay of BL/pGEX4T-1 and BL/OsDOF18 recombinants

Transformed *E. coli* cells were subjected to different abiotic stresses. Induced cultures OD was adjusted to OD<sub>600=1</sub>. Then 10 µL of 50-, 100- and 200-fold diluted bacterial suspension was spotted on LB plates containing 400, 500 and 600 mM NaCl for salt stress; 500, 800 mM and 1M mannitol for dessication. Samples were spotted after 2, 4, 6 and 8 hours of cold stress and after 1, 2 and 3 hours of heat stress.
Plate 18: Spot assay of BL/pGEX4T-1 and BL/OsMYB1 recombinants

Transformed E. coli cells were subjected to different abiotic stresses. Induced cultures OD was adjusted to OD_{600}=1. Then 10 µL of 50-, 100- and 200-fold diluted bacterial suspension was spotted on LB plates containing 400, 500 and 600 mM NaCl for salt stress; 500, 800 mM and 1M mannitol for dessication. Samples were spotted after 2, 4, 6 and 8 hours of cold stress and after 1, 2 and 3 hours of heat stress.
4.8 **IN SILICO CHARACTERIZATION OF OsWRKY71, OsDOF18 AND OsMYB1**

4.8.1 **Intrinsic Disorder in OsWRKY71, OsDOF18 and OsMYB1**

Plant transcription factors are grouped to different transcription factors families, such as MYB, NAC, WRKY and bZIP, have significant degrees of Intrinsic disorder (ID) regions which play a vital role in interactions with DNA and other regulatory proteins (Kragelund et al., 2012). In the past decade, research has revealed that eukaryotic genomes encode 25 to 30% of ID proteins out of which 90% proteins are transcription factors (Liu et al., 2006). Various computational methods have been employed to predict IDRs based on primary protein sequence. These IDs appear to be involved in molecular recognition, for instance, protein binding to DNA to assist transcription or replication.

IDP and IDR can recognize and bind to a multiple partners at different binding sites via short regions, called Molecular Recognition Features (MoRFs) that switch from disorder-to-order confirmation (Uversky, 2011). ID regions enhance the interaction potential of transcription factors which results in a number of protein partners that form a hub (Han et al., 2004) in interactomes. The consequence of ID for protein-protein interactions is now being esteemed in plant science (Kragelund et al., 2012). In plants, NAC transcription factors are the only class of transcription factors in which ID was studied in detail (Jensen et al., 2010).

DISPORED analysis predicted 60% residues to be disordered in OsWRKY71. Only the DNA binding domain region is structured whereas amino acids residues surrounding the DNA binding domain are intrinsically disordered (IDs) (Figure 11A). Structure of DNA binding domain predicted by SMART consists of 59 amino acids (residue 192-250) at the C-terminus (Figure 11B). These flanking disordered regions
Figure 11: Intrinsic disorder (ID) prediction for OsWRKY71

(A) Intrinsic disorder analysis by DISPORED. A cut off value was provided for disorder prediction. A values greater than or equal to 0.05.

(B) Diagrammatic representation of OsWRKY71 structure. OsWRKY71 comprises of a WRKY DNA binding domains.
are supposed to be involved in DNA binding affinity and specificity. But how these ID regions affect the binding is still to be discovered. DISPORED has also predicted five disordered regions in OsWRKY71 involve in protein-protein interaction (Figure 11A).

OsDOF18 contains Zinc-finger DOF domain that make interaction with target gene promoter. The ID prediction profile of OsDOF18 demonstrated that this protein has high degree of IDR regions at both termini. A region at the N-terminus represents putative protein-protein interaction region (Figure 12A). Degree of IDs is 65% in OsDOF18 including a region of high ID of 122 amino acids at C terminus adjacent to DNA binding domain. Only the region of DNA binding domain is structured (Figure 12B).

OsMYB1 contains two MYB DNA binding domains that interact with promoter region of target genes. The graphic organization of OsMYB1 proposed that this protein has numerous regions with IDs, which correspond to putative transcription regulatory domains (TRDs) and protein-protein interaction areas. Four regions were predicted by DISPORED as protein binding regions (Figure 13A). Only the region of DNA binding domain was structured. The location of DNA binding domains is predicted as shown in Figure 13B. Overall degree of ID was about 60% in OsMYB1.

4.8.2 Protein-protein Interaction Network Analysis

To determine whether OsWRKY71, OsDOF18 and OsMYB1 interact with other rice proteins and how they affect the cell function, these proteins were examined for protein-protein interactions by probing the String database. For OsWRKY71, the predicted protein partners included other WRKY proteins for instance, OsWRKY76, OsWRKY53, OsWRKY24 and OsWRKY7, and other proteins including magnesium chetalase, cyclin dependent protein kinase, and ZIM domain containing proteins
Figure 12: Intrinsic disorder (ID) prediction for OsDOF18

(A) Intrinsic disorder analysis by DISPORED. A cut off value was provided for disorder prediction i.e. values greater than or equal to 0.05
(B) Diagrammatic representation of OsDOF18 structure. OsDOF18 contains a zinc finger DOF DNA binding domains.
Figure 13: Intrinsic disorder (ID) prediction for *OsMYB1*

(A) Intrinsic disorder analysis by DISPORED. A cut off value was provided for disorder prediction. Values greater than or equal to 0.05 (black bar).

(B) Diagrammatic representation of *OsMYB1* structure. *OsMYB1* comprises of 2 SANT DNA binding domains.
(Figure 14). For OsDOF18, only one interaction partner was predicted which is AP2 domain containing protein (Figure 15). OsMYB1 was predicted to bind ATP-dependent RNA helicase, TAZ zinc finger proteins and phosphoribosylamine glycine ligase (Figure 16). Predicted protein partners have been found to be involved in various biological processes, for instance, photosynthesis, signal transduction and also in providing defense against various stresses.

Conclusively, present study highlights the in silico identification of stress responsive cis-regulatory elements in the promoter of OsRGLP2 gene and their corresponding binding proteins. The DNA binding domains of the identified proteins OsWRKY71, OsMYB1 and OsDOF18 were able to interact with OsRGLP2 promoter as confirmed by EMSA. OsWRKY71 is characterized for its role in biotic stresses but no work has been done for its role in response to abiotic stresses. OsMYB1 and OsDOF18 are the proteins, yet to be characterized in rice. OsWRKY71, OsDOF18 and OsMYB1 transcript level was increased in response to various abiotic stresses suggests the importance of these proteins in providing protection to rice under abiotic stresses. Availability of more information regarding these proteins in rice may further enhance our understanding about gene expression and its manipulation during biotic as well as abiotic stresses.
Figure 14: Protein-protein interaction network of OsWRKY71 analyzed by String

(A) Interaction network of OsWRKY71 with differentially expressed proteins. In the resulting proteins network, proteins are represented by nodes which are linked by lines whose thickness tells the confidence level. Various line colors denote different kinds of associations: red line represents fusion association evidence; green line represents neighborhood association; light blue line indicates database evidence; black line represents coexpression evidence; purple line denotes experimental evidence of proteins association; yellow line represents text-mining evidence and blue line indicates co-occurrence. (B) Coexpressed proteins evidence from O. sativa and other organisms.
Figure 15: Protein-protein interaction network of OsDOF18 analyzed by String

(A) Interaction network of OsDOF18 with differentially expressed proteins. Black line shows the coexpression evidence (B) Coexpressed proteins evidence from O. sativa only.
Figure 16: Protein-protein interaction network of OsMYB1 analyzed by String

(A) Interaction network of OsMYB1 with differentially expressed proteins. In the resulting proteins network, proteins are represented by nodes which are linked by lines whose thickness tells the confidence level. Various line colors denote different kinds of associations: red line represents fusion association evidence; green line represents neighborhood association; light blue line indicates database evidence; black line represents coexpression evidence; purple line denotes experimental evidence of proteins association; yellow line represents text-mining evidence and blue line indicates co-occurrence (B) Coexpressed proteins evidence from O. sativa and other organisms.
SUMMARY

Plants are immobile organisms and frequently subjected to diverse environmental stresses, including drought, salt, and threshold temperatures. These abiotic stresses create a severe risk to crop plants growth and productivity. Plants have developed different adaptive strategies to lessen the unfavorable consequences by changing their molecular and cellular functions, e.g altering the gene expression and consequent action of their gene products. Germin like proteins (GLPs) are cell wall proteins and their involvement in different stresses and other developmental pathways have been established. In addition to other approaches, functional imminent into the GLPs expression can be attained by performing the analysis of their regulatory regions and respective binding proteins i.e transcription factors. Earlier Mahmood (2007) isolated, cloned and characterized the OsRGLP2 promoter region. Heterologous expression of reporter genes under OsRGLP2 promoter confirmed that this promoter directed the GUS overexpression in wounding, salinity and drought stresses.

The present study was designed to isolate and characterize the proteins that bind with the stress responsive cis-acting regulatory elements of OsRGLP2 promoter. For this purpose, OsRGLP2 promoter sequence was analyzed with PLACE, PLANTCARE, ConSite and JASPAR online databases to identify stress responsive regulatory motifs located on the promoter sequence. The OsRGLP2 promoter region was enriched with W-Box, AAAG and WAACCA motifs. Suggested proteins for these regulatory motifs were then subjected to BLAST with Rice genome to identify the target proteins from Rice genome. OsWRKY71, OsMYB1 and OsDOF18 were chosen as targets for these regulatory motifs.

Bioinformatics analysis was performed to identify the DNA binding domains and characteristics of OsWRKY71, OsMYB1 and OsDOF18. DNA binding
domains of OsWRKY71, OsMYB1 and OsDOF18 with overhangs of about 15 amino acids were cloned in pGEX4T-1. Recombinant vectors were overexpressed in *E. coli* and recombinant proteins were purified by using various high throughput purification techniques. To check the interaction of recombinant proteins with *OsRGLP2* promoter, electrophoretic mobility shift assay (EMSA) was carried out. Various probes were designed on different regions of *OsRGLP2* promoters containing regulatory motifs for *OsWRKY71, OsMYB1* and *OsDOF18*. GST-*OsWRKY71* was able to bind to *OsRGLP2* promoter fragments containing either CTGACC, TTGACT or TGAC core sequence. Additionally, EMSA analysis has confirmed the binding of GST-*OsMYB1* and GST-*OsDOF18* with promoter fragments containing WAACCA and AAAG motifs respectively. Binding was further confirmed by competitor EMSA and EMSA with mutant oligonucleotides. Shifts were also observed with all the probes when EMSA was carried out with nuclear protein extract from rice.

To examine the expression of *OsWRKY71, OsMYB1* and *OsDOF18* under various abiotic stress conditions in rice, available microarray data from GENEVESTIGATOR was used. Analyses of microarray data demonstrated that expression of *OsWRKY71, OsMYB1* and *OsDOF18* genes is regulated by various abiotic stress conditions. To validate this expression analysis, Quantitative Real-Time PCR analysis was carried out. *OsWRKY71* showed up-regulation in response to cold and salt stresses while no significant change was observed in response to wounding, heat and drought stresses. Transcript level of *OsDOF18* was up-regulated in case of salt and drought stresses. *OsMYB1* expression was 23 fold higher in response to wounding which suggested the importance of inducible *OsMYB1* in wound signaling in rice. There was no significant change in the expression of *OsMYB1* in case of salt, drought, heat and cold treatments.
In order to examine the function of proteins against diverse abiotic stresses, recombinant plasmids were expressed in *E. coli*. The differential response of *E. coli* harbouring pGEX-*Os*WRKY71, pGEX-*Os*DOF18 and pGEX-*Os*MYB1 was inspected for their diverse level of tolerance to various abiotic stresses in comparison to empty pGEX-4T1 vector. In conclusion, the positive response of *Os*WRKY71, *Os*DOF18 and *Os*MYB1 in abiotic stresses suggests their association with *Os*RGLP2 promoter. The results demonstrated that *Os*WRKY71, *Os*MYB1 and *Os*DOF18 are stress related genes in rice and overexpression of *Os*WRKY71, *Os*DOF18 and *Os*MYB1 genes in crop plants may help in obtaining stress tolerant lines.


Diaz, I., J. Vicente-Carbajosa, Z. Abraham, M. Martinez, M. I. Isabel-La and P. Carbonero. 2002. The GAMYB protein from barley interacts with the DOF


Lopez-Molina, L., S. Mongrand and N. H. Chua. 2001. A post germination developmental arrest checkpoint is mediated by abscisic acid and requires the


Manosalva, P. M., R. M. Davidson, B. Liu, X. Zhu, S. H. Hulbert, H. Leung and J. E.


polymerase with open complex formation at PRM is relieved by a 10-base pair deletion between the two promoters. J. Biol. Chem., 270(51): 30428-30433.


Sakuma, Y., Q. Liu, J. G. Dubouzet, H. Abe, K. Shinozaki and K. Yamaguchi-


and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. Plant Cell, 16(9): 2481-2498.


Whitson, P. A. and K. S. Matthews. 1986. Dissociation of the lactose repressor-


APPENDICES

Appendix 1

Half Strength MS Media

MS salts 2.22g/L

Plant Agar 0.7%

Sucrose 15 g/L

pH was adjusted to 5.8, autoclaved at 121°C, 15 psi for 15 minutes.

Appendix 2

Gel formulations (10 ml)

<table>
<thead>
<tr>
<th>Gel percent</th>
<th>Deionized Water (mL)</th>
<th>30% acryl/bis (mL)</th>
<th>Gel buffer* (mL)</th>
<th>10% SDS w/v (mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4%</td>
<td>6.1</td>
<td>1.3</td>
<td>2.5</td>
<td>0.1</td>
</tr>
<tr>
<td>12%</td>
<td>3.4</td>
<td>4</td>
<td>2.5</td>
<td>0.1</td>
</tr>
<tr>
<td>15%</td>
<td>2.4</td>
<td>5</td>
<td>2.5</td>
<td>0.1</td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Separating Gel Buffer-1.5 M Tris-HCl pH 8.8
* Stacking Gel Buffer-0.5 M Tris-HCl pH 6.8

1.5 M Tris-HCl, pH 8.8 (Separating buffer)

Tris base 18.15 g

Distilled water 100 ml

Add some deionized water mix and then bring volume upto 100ml. Adjust to pH 8.8 with 6N HCl. Store at 4°C.
0.5 M Tris-HCl, pH 6.8 (Stacking buffer)
Tris Base 6.0 g
Add some deionized water, mix and then bring volume up to 100 ml. Adjust to pH 6.8 with 6N HCl. Store at 4°C.

Appendix 3
10x Electrode (running) buffer, pH 8.3
Tris base 30.3 g
Glycine 144 g
SDS 10 g
Add deionized water, bring volume up to 1L. Store at 4°C. Dilute 100 ml 10x stock with 900 ml deionized water for each electrophoretic run.

Appendix 4
Elution Buffer
Tris-HCl 50 mM pH 8.0
Reduced glutathione 10 mM

Appendix 5
High pH buffer
Tris-HCl 0.1 M pH 8.5
NaCl 0.5 M

Appendix 6
Low pH Buffer
Na-acetate 0.1 M pH 4.5
NaCl 0.5 M

Appendix 7
TEN Buffer
Tris-HCl 10 mM pH 8
EDTA 1 mM
NaCl 0.1 M

Appendix 8
Wash Buffer
Maleic acid 0.1 M pH 7.5
NaCl 0.15M
Tween 20 0.3%

Appendix 9
Blocking Solution (10X)
Blocking reagent 10 g
Maleic acid buffer 100 mL
Dissolve well and autoclave.

Appendix 10
Detection buffer
Tris-HCl 0.1 M pH 9.5
NaCl 0.1 M

Appendix 11
6% Native EMSA gel
Acryl amid:Bis (29:1) 4 mL
TBE (5x) 4 mL
H$_2$O 14 mL
TEMED 40 μL
10% APS 80 μL

Appendix 12
SSC Buffer (2X)
NaCl 300 mM
Na-citrate 30mM
Appendix 13

Nuclear Proteins Extraction Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCl</td>
<td>50 mM</td>
</tr>
<tr>
<td>Sucrose</td>
<td>0.3 M</td>
</tr>
<tr>
<td>KCl</td>
<td>15 mM</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>5 mM</td>
</tr>
<tr>
<td>EDTA</td>
<td>0.1 mM</td>
</tr>
<tr>
<td>DTT</td>
<td>0.1 mM</td>
</tr>
</tbody>
</table>

pH 7.5

Appendix 14

Protein Isolation Buffer

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCl</td>
<td>10 mM</td>
</tr>
<tr>
<td>DTT</td>
<td>1 mM</td>
</tr>
<tr>
<td>NaCl</td>
<td>0.4 M</td>
</tr>
</tbody>
</table>

pH 7.5