Cloning, expression and characterization of Aquaporins from Wheat in insect cells using a baculovirus expression system

A thesis submitted for the award of
Masters of Biotechnology (by research)

Rajeswari Vaduganathan

Faculty of Life and Social Sciences (FLSS)
Swinburne University of Technology
Australia

2013
The final analysis of the life you led on this planet is between you and God, not between you and others.

- Mother Teresa.
Abstract

Aquaporins are water transport proteins that possess important roles in plant-water relationship. The allopolyploid nature of wheat makes this staple cereal crop more difficult to sequence and understand the diversity of aquaporin proteins. Most of the aquaporins in rice, Arabidopsis and maize shows tissue specific expression. This study aimed to identify, express and characterize root specific aquaporin proteins of wheat and investigate their specific functions.

Two full length cDNAs of root specific aquaporins from common wheat (*T.aestivum* L. cv. Cranbrook) *Triticum aestivum* were isolated. The newly isolated sequences were identified as a PIP2 and PIP1 aquaporin family members by comparison with the Rice and existing wheat aquaporin genes in the database. Additionally one possible non-functional (alternate spliced) PIP2 aquaporin was isolated.

This study found that the PIP2 gene was highly expressed at the early seedling stage (8 day old plant) of the plants than at the heading and flowering stage. This correlates with the higher water demand for the growing plant at early stages.

In this study the insect cell/baculovirus system was used to express the identified full length PIP1 and PIP2 genes. The functional analysis of the wheat and human AQP proteins was performed by subjecting the whole insect cell to hypo-osmotic conditions and the size of the whole insect cell was measured using ImageJ software. Results indicate that the water treatment for the infected and uninfected insect cells was unsuccessful because of the technological problems and time constraints. Further glicerol transport and mercury sensitive aquaporins can be tested using the proposed method. This will be the first plant aquaporin expressed in the baculovirus system to demonstrate water transport activity and substrate specificity.
Acknowledgements

I would like to thank my supervisors Dr. Tony Barton and Professor Mrinal Bhave, without whom this research would not be possible. I am grateful to thank Tony for his constant support and advice for the experiments, and for sacrificing so much of his time for long meetings and troubleshooting. I also thank Mrinal for her suggestions in the project and thesis chapters.

Many Thanks to laboratory technicians Ngan Nguyen, Alison Demaio, Sheila Curtis, Savithri Galappathie, Soula Mougos and Chris Keys for their timely help and technical assistance throughout the project.

Thanks to Edmond Linossi, Ravi Nirmal and Neeti for helping me in few experiments.

In addition I would like to extend my thanks to the past and present PhD students Natallia, Abdullah, Shahani, Babu, Kerrie, Mark, Huimei, Pete, Shee Ping, Abi, Bita, Jacqui B, Khanh, Hayden, Hamid, Suchetana, Jun, Jiaway, Kristin, Monir, Vanu, Gaurav, Anjula, Lalitha, Aravind, Dhivya and Swarna for all their support and advice throughout the project.

This thesis would not be possible without the love and support of my beloved family members, mom, dad, sisters, brother-in-law and nephews. I sincerely thank every one of them for their constant support at every stage.

Rajeswari Vaduganathan
Declaration

I, Rajeswari Vaduganathan, declare that the Master of Biotechnology (by research) thesis entitled –Expression and Characterization of Aquaporins from wheat in Baculovirus” contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma, and has not been previously published by another person. Except where otherwise indicated, this thesis is my own work.

Rajeswari Vaduganathan
October 2013
# Table of Contents

1 Literature review

1.1 Introduction 2

1.1.1 Wheat and its importance as a cereal crop 2

1.1.2 Water transport in plants 2

1.2 Aquaporins 4

1.2.1 Historical background of aquaporin research 5

1.3 Aquaporin structure 6

1.3.1 Typical structure of an aquaporin monomer 6

1.3.2 Aquaporin tetramers 7

1.4 Plant aquaporins and their molecular diversity 8

1.4.1 PIPs: Plasma Membrane Intrinsic Proteins 10

1.4.2 TIPs: Characteristic vacuolar membrane aquaporins 10

1.4.3 SIPs (Short basic intrinsic proteins) 11

1.4.4 NIPs (NOD26-like intrinsic proteins) 11

1.5 Conserved residues of PIP aquaporin Family 12

1.6 Spinach aquaporins 13

1.6.1 Water transport of spinach PIPs and gating of plant AQP 13

1.7 Rice aquaporins 16

1.7.1 Localization of Expression of ten OsPIP genes in rice seedlings 16

1.8 Wheat aquaporins 17

1.8.1 Water transport of wheat PIPs 17

1.8.2 Regulation and Physiological roles of PIPs in *Triticum aestivum*s 18

1.9 Other functions of aquaporins 20

1.9.1 Aquaporins as CO₂ transporters 20

1.9.2 Aquaporins as H₂O₂ transporters 21

1.10 Heterologous expression systems 21

1.10.1 Introduction 21

1.10.2 Bacterial expression system 22

1.10.2.1 Expression in *E.coli* 22
Table of Contents

1.10.3 Expression in yeast cells 24
1.10.4 Expression in insect cells 24
   1.10.4.1 Baculovirus expression system 24
1.10.6 Problems in heterologous expression system 25

1.11 Functional testing of aquaporins expressed in heterologous expression system 26
   1.11.1 Functional testing of aquaporins in Yeast 26
      1.11.1.1 Review of Light scattering 26
      1.11.1.2 Functional testing of aquaporins in Yeast using Stopped-flow spectrometry 27
      1.11.1.3 Ammonia transport in the yeast cell 27
   1.11.3 Functional testing of aquaporins in *Xenopus laevis* 28
      1.11.3.1 Water permeability in *Xenopus oocytes* 29

1.12 Recombinant expression of membrane proteins 29
   1.12.1 Factors necessary for the production of the functional eukaryotic membrane protein 29
   1.12.2 Multistep process in expressing a functional membrane protein 30

1.13 Regulation of aquaporin 30
   1.13.1 Regulation of plant aquaporin by phosphorylation and membrane trafficking 30
   1.13.2 Glycosylation of Aquaporins 31

1.14 Aquaporin methylation 32

1.15 Protein purification 32

1.16 Aims of this project 33

1.17 Thesis structure 34

2 Cloning and sequencing of two *TaPIP* genes at both genomic and cDNA level 37

2.1 Introduction 37
   2.1.1 PIPs 37
   2.1.2 Criteria for gene selection 37
2.2 Materials and Methods

2.2.1 Prepared solutions

2.2.1.1 Buffers and solutions

2.2.1.2 Sterilisation

2.2.2 Plant material

2.2.3 Sterilization of seeds

2.2.4 Plants growth condition

2.2.5 Bioinformatics Methods

2.2.5.1 *In silico* search methods to identify Transcript Assemblies for PIP1 and PIP2

2.2.5.2 *In silico* search methods to identify expression pattern of rice aquaporins

2.2.6 Primer design

2.2.6.1 Design of primers to amplify full length PIP genes from wheat

2.2.7 Molecular methods

2.2.7.1 Genomic DNA purification

2.2.7.2 Quantification of DNA

2.2.7.2.1 Spectrophotometric quantification

2.2.7.2.2 Agarose gel electrophoresis

2.2.8 Polymerase Chain Reaction (PCR)

2.2.8.1 PCR conditions

2.2.8.2 Primer synthesis

2.2.8.3 Amplification of full length *PIP* genes of common wheat

2.2.9 Gel purification

2.2.9.1 DNA bands purified from agarose gel

2.2.10 Cloning of PCR products

2.2.10.1 Ligation

2.2.10.2 Transformation

2.2.11 Colony PCR

2.2.12 Plasmid DNA purification
Table of Contents

2.2.13 Restriction digestions 49
2.2.14 DNA sequencing and sequence analysis 49
  2.2.14.1 Purified plasmid DNA for DNA sequencing 49
  2.2.14.2 DNA sequence analysis 50
  2.2.14.3 DNA sequencing of clones of PIP genes from genomic DNA and cDNA 50
2.2.15 RNA Purification 50
  2.2.15.1 Removal of RNase from materials 50
  2.2.15.2 Total RNA extraction from wheat shoot (as well as from leaves earlier) and root tissues 51
  2.2.15.3 DNase treatment of total RNA 51
  2.2.15.4 Agarose gel electrophoresis of RNA samples 51
  2.2.15.5 Spectrophotometric Quantification of RNA 51
2.2.16 cDNA synthesis 51
2.2.17 cDNA quality assessment 52
2.2.18 Cloning of full length PIP cDNA sequences 53
2.2.19 Cloning of full length PIP genomic DNA sequences 54
2.2.20 Analysis of wheat PIP gene sequences 54
2.2.21 Analysis of the protein structures and features of wheat PIP genes 55
2.2.22 Identification of the Orthologous rice and Barley gene for TaPIP1 and TaPIP2 55
2.2.23 Analysis of gene structure of the full length wheat AQP genes TaPIP1 and TaPIP2 55

2.3 Results and Discussion 56
  2.3.1 Attempts to amplify full length TaPIP2 and TaPIP1 gene 56
  2.3.2 cDNA quality testing 57
  2.3.3 Isolation of two full length aquaporin genes from wheat 58
    2.3.3.1 Isolation of full length TaPIP2 gene from wheat 60
      2.3.3.1.1 Amplification of TaPIP2 gene from genomic DNA 60
      2.3.3.1.2 Subgrouping of TaPIP2 gene from 61
Table of Contents

2.3.3.1.3 Amplification of TaPIP2 gene from cDNA 64
2.3.3.1.4 Analysis of full length cds and alternate splice variant of TaPIP2 gene 65
2.3.3.1.5 Comparison of TaPIP2 gene from root cDNA with the predicted cds of TaPIP2-9a and TaPIP2-9b obtained from genomic DNA 67
2.3.3.1.6 Gene structure of TaPIP2-9 gene 67
2.3.3.2 Isolation of full length TaPIP1 gene from wheat 70
2.3.3.2.1 Amplification of TaPIP1 gene from cDNA 70
2.3.3.2.2 Subgrouping of TaPIP1 gene from genomic DNA 72
2.3.3.2.3 Amplification of TaPIP1 gene from cDNA 79
2.3.3.2.4 Comparison of TaPIP1 gene from root cDNA with the TaPIP1-6a and TaPIP1-2a obtained from the genomic DNA 80
2.3.3.2.5 Gene structure of TaPIP1-2a and TaPIP1-6a gene 80
2.3.4 Comparison of the TaPIP1-2a and TaPIP2-9a coding sequence with the known full length ESTs 82
2.3.5 Relationship between the TaPIP2-9a cds and TaPIP1-2a cds is orthologous to Barley and Rice 84
2.3.5.1 The cDNA of TaPIP2-9a gene 84
2.3.5.2 The gDNA of TaPIP2-9a gene 85
2.3.5.3 The cDNA TaPIP1-2a gene 86
2.3.5.4 The gDNA of TaPIP1 gene 86
2.3.6 Prediction of substrate specificity and Mercury sensitivity of wheat AQP proteins 88
2.3.7 Prediction of ammonia transport sites in the wheat AQP 90
Table of Contents

proteins
2.3.8 Prediction of important amino acid residues in wheat aquaporins compared to rice 90
2.3.9 Candidate aquaporins to facilitate water uptake in barley 91
2.4 Summary and future work 95

3 Heterologous expression of aquaporin genes from human and wheat in Baculovirus infected insect cells.

3.1 Introduction 98
3.1.1 Phase separation purification of integral membrane proteins 99
3.1.2 La Protein 99
3.2 Methods 100
3.2.1 Sf21 cell culture 100
3.2.2 Co-transfection 100
3.2.3 PCR verification of recombinant virus 101
3.2.3.1 Viral DNA extraction by phenol chloroform method 101
3.2.3.2 Viral DNA extraction by isopropanol method 102
3.2.3.3 PCR verification of recombinant baculoviruses 103
3.2.4 Expansion of recombinant baculoviruses 104
3.2.5 Infection of recombinant baculoviruses 104
3.2.6 Total Membrane protein extraction 104
3.2.7 Homogenization using a syringe and needle 104
3.2.8 Sorvall centrifugation 104
3.2.9 Bradford assay 104
3.2.10 Phase Separation using detergent Triton X-114 105
3.2.11 Affinity Chromatography using IMAC (Immobilized Metal ion Affinity Chromatography) 106
3.2.12 SDS PAGE 106
3.2.13 Analysis of the levels of Aquaporins by Western blotting 107
3.2.14 Detection of western blot with DAB 108
3.3 Result and Discussion 108
3.3.1 Preparation of pBacPAK9 and pBacPAK His1 plasmid DNA 108
Table of Contents

3.3.2 Directional cloning of AQPs into Histidine tagged and untagged baculoviral expression vectors 109
  3.3.2.1 Directional cloning of wheat aquaporins into tagged and untagged baculoviral expression vectors 109
  3.3.2.2 Directional cloning of mammalian AQP3 into tagged and untagged baculoviral expression vectors 111
3.3.3 Verification of the Recombinant baculovirus 113
  3.3.3.1 PCR on Viral DNA 113
  3.3.3.2 PCR using Baculoviral Vector based primers 114
  3.3.3.3 PCR using Sequence specific primers 114
3.3.4 Expression of AQPs 116
  3.3.4.1 Phase separation 116
  3.3.4.2 Expression of wheat and mammalian aquaporin at 48, 72 hr infection and from frozen cell pellet 118
  3.3.4.3 Concentration of the proteins 121
3.3.5 Immunoblot analysis 121
  3.3.5.1 Immunoblot analysis of hAQP3 121
  3.3.5.2 Immunoblot analysis of wheat AQP 121
3.4 Summary 127

4 Functional testing of aquaporins expressed in insect cells

4.1 Introduction 130
  4.1.1 Glycerol transport 130
  4.1.2 Role of glycerol in mammals and plants 130
  4.1.3 Plant Aquaglyceroporins 131
  4.1.4 Aquaglyceroporins 131
  4.1.5 Existing methods to test the function of the aquaporins: advantages and disadvantages 132
  4.1.6 Proposed method from this study: advantages and disadvantages 133
  4.1.7 Cell volume as a measure of water transport 133

xvi
#### Table of Contents

4.2 Materials and Methods 134  
4.2.1 Experimental procedures 134  
4.2.2 Method development trail 134  
4.2.2.1 Water, Glycerol and Mercury experiment triplicates 134  
4.2.2.2 Image processing 135  
4.2.2.2.1 Measuring the insect cells size using ImageJ 135  
4.2.3 MATLAB 136  
4.3 Results and Discussion 137  
4.3.1 Hypo-osmotic water transport and mercury sensitivity 137  
4.3.2 Isoosmotic glycerol transport 141  
4.3.2 Statistical analysis 142  
4.4 Summary 143

5 **General conclusions and future direction**  
5.1 Conclusions and contribution from this work 145  
5.2 Future direction 147

6 **References** 150

7 **Appendices** 169
List of Figures

Figure 1.1 Water transport pathway in the roots of a plant. 3
Figure 1.2 The membrane topology of human AQP1. 6
Figure 1.3 The hourglass model for the folding of human AQP1 8
Figure 1.4 Model of PM28A (spinach leaf plasma membrane) and mechanism of plant AQP gating by blocking the water pore. 15
Figure 1.5 Transport of ambient CO₂ to the chloroplast through several barriers. 20
Figure 1.6 Agre’s experiment with *Xenopus oocyte* cells containing or lacking aquaporin. 28

Figure 2.1 Example of quality test of cDNA preparations 58
Figure 2.2 PCR with gDNA of *TaPIP2* gene with different primer pairs. 61
Figure 2.3 Wheat *TaPIP2* gDNA sequences. 63
Figure 2.4 Amplification of full length *TaPIP2* gene from 8–day-old root cDNA. 64
Figure 2.5 Primer positions of *TaPIP2-9* gene. 65
Figure 2.6 Wheat PIP2 cds sequences 66
Figure 2.7 Gene structure of *TaPIP2-9* gene. 68
Figure 2.8 PCR to amplify full length *TaPIP1* gene from gDNA, gel purified and restriction digested product. 71
Figure 2.9 Alignment of *TaPIP1-2a* with *TaPIP1-2* (EU177547) and *TaPIP1-6a* and *TaPIP1-6* (EU177551) genomic sequence. 78
Figure 2.10 PCR to amplify full length *TaPIP1* gene from shoot and root cDNA. 79
Figure 2.11 Gene structure of *TaPIP1* gene 81
Figure 2.12 The Phylogenetic tree (ClustalW alignment of deduced amino acid sequence) demonstrating orthology between rice, barley and *A.thaliana* PIPs. 87
Figure 2.13 Kyte-Doolittle plots of Wheat aquaporins 88
Figure 2.14 Phylogenetic tree (ClustalW alignment of deduced amino acid sequence) demonstrating the substrate specificity of AQPs. 89
Figure 2. Alignment of glycerol, water transport and Mercury sensitive AQPs.

Figure 3.1 Construction of baculovirus recombinants with BacMagic systems by homologous recombination.

Figure 3.2 Putative amino acid sequences of Histidine tagged wheat aquaporins.

Figure 3.3 Plasmid DNA analysis of the pBacPAK9 constructs of native wheat aquaporins by restriction digestion.

Figure 3.4 Plasmid DNA analysis of the pBacPAKHis1 constructs of native wheat aquaporins by restriction digestion.

Figure 3.5 Putative amino acid sequence of Histidine tagged hAQP3.

Figure 3.6 Restriction digests confirmation of transfer vector plasmid DNA of untagged and tagged hAQP3.

Figure 3.7 PCR on Viral DNA to Screen for the His tagged (pBacPAKHis1-TaPIP2-9a and pBacPAKHis1-TaPIP1-2a) and untagged (pBacPAK9-TaPIP2-9a and pBacPAK9-TaPIP1-2a) wheat aquaporin Gene Cassette.

Figure 3.8 PCR on Viral DNA to Screen for the His tagged (pBacPAKHis1-hAQP3) and untagged (pBacPAK9-hAQP3) mammalian aquaporin Gene Cassette using Phenol Chloroform method (A) and Isopropanol method (B).

Figure 3.9 Phase separation of hydrophilic and amphiphilic proteins.

Figure 3.10 SDS PAGE and Western blot analyses of His tagged wheat AQPs, TaPIP2-9a and TaPIP1-2a.

Figure 3.11 Concentration of His tagged hAQP3 from 200 mL culture without and with 5% Imidazole verified by a Coomassie-stained gel and immunoblot analysis of His tagged hAQP3.

Figure 3.12 Concentration of His tagged TaPIP2-9a and His tagged TaPIP1-2a from 200 mL culture of 72 h infected insect cells, verified by a Coomassie-stained gel and immunoblot analysis of His tagged hAQP3.
List of Figures

Figure 4.1  Volume changes of insect cells in response to hypo-osmotic shock. 140
Figure 4.2  Volume changes of insect cells in response to isoosmotic shock. 141
List of Tables

Table 1.1 Organ specific expression rice aquaporin genes. ......................................................................................................................... 16
Table 1.2 Genes isolated from gDNA of wheat and the corresponding accession number. .................................................................................. 19
Table 2.1 Consensus residues distinguishing AQP salts and GLP proteins. .................................................................................................................. 39
Table 2.2 Wheat cultivar used in the study. ......................................................................................................................................................... 40
Table 2.3 Primers used to synthesize cDNA. .................................................................................................................................................... 52
Table 2.4 Primers used to assess the quality of cDNA. .............................................................................................................................. 53
Table 2.5 Primers designed to amplify full length wheat PIP genes for this study. ............................................................................................ 57
Table 2.6 TA (Transcript assembly) matches obtained for the partial TaPI1-5 (EU177550) and TaPI2-8 (EU177561) genes. ................................................. 59
Table 2.7 Summary of the % identities of full length TaPI2-9 gene at genomic level with the partial TaPI genes from the database. ......................... 62
Table 2.8 % identity of introns and exons of TaPI2-9a gene isolated from gDNA and cDNA of wheat. ................................................................. 69
Table 2.9 Summary of the % identities of full length TaPI1-5 gene at genomic level with the partial TaPI genes from the database. ......................... 72
Table 2.10 % identity of introns and exons of TaPI1-5 gene isolated from gDNA of wheat. ................................................................. 81
Table 2.11 % identity of genes isolated from gDNA and cDNA to EST. ........................................................................................................ 83
Table 3.1 Primers used to verify the Viral DNA insert. .......................................................................................................................... 103
Table 4.1 Aquaglyceroporins in mammals and plants. .......................................................................................................................... 134
Appendices

Appendix 1  Alignment of *TaPIP1-2a* full length coding sequence (cds), *TaPIP1-2a* and *TaPIP1-2* (gDNA) to predict intron and exon junction. 170

Appendix 2  Alignment of *TaPIP2-9* gene obtained from cDNA and genomic DNA to assign intron and exon. 173

Appendix 3  ClustalW alignment of deduced amino acid sequence of *TaPIP1-2a cds* and *TaPIP2-9a cds* with Rice MIPs. 176

Appendix 4  Accession numbers of deduced amino acid sequence of *H.vulgare, O.sataiva, A.thaliana, T.aestivum* and *T.turgidum*. 180

Appendix 5  Expression pattern of Rice *OsPIP1-2* putative expressed protein. 181

Appendix 6  ) Alignment of predicted amino acid sequence of *TaPIP2-9a cds* obtained from root cDNA with all the known complete cds of wheat AQPs in the database. 182
  
  ) Alignment of predicted amino acid sequence of *TaPIP1-2a cds* obtained from root cDNA with all the known complete cds of wheat AQPs in the database. 183

Appendix 7  Uninfected insect cells before and after 72hrs incubation. 184

Appendix 8  Procedure for automated ImageJ software 185
**Abbreviation**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg (or R)</td>
<td>arginine</td>
</tr>
<tr>
<td>At</td>
<td><em>Arabidopsis thaliana</em></td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>BDT</td>
<td>Big Dye Terminator</td>
</tr>
<tr>
<td>cDNA</td>
<td>copy DNA</td>
</tr>
<tr>
<td>CDS/cds</td>
<td>coding sequence</td>
</tr>
<tr>
<td>cRNA</td>
<td>complementary RNA</td>
</tr>
<tr>
<td>C-terminal</td>
<td>carboxyl terminal</td>
</tr>
<tr>
<td>cv</td>
<td>cultivar</td>
</tr>
<tr>
<td>Cys (or C)</td>
<td>cysteine</td>
</tr>
<tr>
<td>dH₂O</td>
<td>deionised water</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>double deionised water</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNase</td>
<td>deoxyribonuclease</td>
</tr>
<tr>
<td>d.p.i</td>
<td>days post infection</td>
</tr>
<tr>
<td>EST</td>
<td>expressed sequence tag</td>
</tr>
<tr>
<td>gDNA</td>
<td>genomic DNA</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>Gln (or Q)</td>
<td>glutamine</td>
</tr>
<tr>
<td>Hv</td>
<td>Hordeum vulgare (barley)</td>
</tr>
<tr>
<td>Ile (I)</td>
<td>isoleucine</td>
</tr>
<tr>
<td>Kb</td>
<td>kilobase pairs</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodaltons</td>
</tr>
<tr>
<td>Leu (or L)</td>
<td>leucine</td>
</tr>
<tr>
<td>Met (or M)</td>
<td>methionine</td>
</tr>
<tr>
<td>M.O.I</td>
<td>Multiplicity of Infection</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NIP</td>
<td>nodulin26-like membrane intrinsic protein</td>
</tr>
<tr>
<td>NPA</td>
<td>asparagine-proline-alanine motif</td>
</tr>
<tr>
<td>N-terminal</td>
<td>amino terminal</td>
</tr>
<tr>
<td>Os</td>
<td>Oryza sativa (rice)</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>Phe (or F)</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>PIP</td>
<td>Plasma membrane Intrinsic Protein</td>
</tr>
<tr>
<td>PM</td>
<td>plasma membrane</td>
</tr>
<tr>
<td>Pro (or P)</td>
<td>proline</td>
</tr>
<tr>
<td>RE</td>
<td>restriction enzyme</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>Ser (or S)</td>
<td>serine</td>
</tr>
<tr>
<td>Sf</td>
<td><em>Spodoptera frugiperda</em></td>
</tr>
<tr>
<td>SIP</td>
<td>Small, basic membrane Intrinsic Protein</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>Ta</td>
<td>Triticum aestivum (common wheat)</td>
</tr>
<tr>
<td>TE</td>
<td>tris-EDTA buffer</td>
</tr>
<tr>
<td>Thr (or T)</td>
<td>threonine</td>
</tr>
<tr>
<td>TIP</td>
<td>Tonoplast Intrinsic Protein</td>
</tr>
<tr>
<td>TMH</td>
<td>transmembrane helical domain</td>
</tr>
<tr>
<td>Tyr (or Y)</td>
<td>tyrosine</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>UIC</td>
<td>Uninfected insect cell</td>
</tr>
<tr>
<td>Val (or V)</td>
<td>valine</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside</td>
</tr>
</tbody>
</table>
1

Literature Review
1 Literature review

1.1 Introduction

1.1.1 Wheat and its importance as a cereal crop

Wheat (Triticum aestivum L.), is a monocot which belongs to the Poaceae (grass family). Wheat is one of the most important staple food crops of the world. Wheat was one of the first crops to be domesticated 10,000 years ago and is adapted to temperate regions of the world. Common wheat with a large genome (16,000 Mb), is comprised of three subgenomes A, B and D and has a high proportion (~80%) of repetitive sequences. Therefore this cereal crop has been a difficult genome to analyze. Each sub genome has 7 chromosomes, making \( n = 21 \). This temperate cereal crop occupies one sixth (17%) of crop acreage worldwide, feeding nearly half (40%) of the world population and providing one fifth (20%) of the total food calories and protein in human nutrition. Wheat productivity is limited by lack of water essential for growth (Rebetzke et al., 2008).

1.1.2 Water Transport in Plants

Most plants secure the water they need from the roots. In roots, water flows radially from the root surface to the central cylinder. Soil water enters the root through its epidermis and travels in both the cytoplasm of the root cell called symplast (water crosses the plasma membrane and then passes from cell to cell through plasmodesmata). Symplast is the major pathway for movement of water through the cells of the root and the nonliving parts of the root called apoplast (water in the spaces between the cells and in the cell walls themselves, this water has not crossed a plasma membrane), and in the apparent free space of the roots and the xylem vessels, where dissolved substances diffuse freely. In roots the apoplast is discontinuous due to the presence of a ring of the endodermis with thick casparian strips (a band of suberized cell walls in the radial and transverse walls of the endodermis). The water then passes into the cortex and reaches the endodermis (Fig 1.1). The thick walls of endodermal cells will not allow the water to diffuse due to the
casparian strip. Therefore to enter the center of the root (stele), apoplastic water must enter the symplasm of the endodermal cells and from here it can pass by plasmodesmata into the center of the root cells (stele). From the stele, water is free to move between and through the cells. Water directly enters into the xylem of the young roots. In older roots water passes into the phloem, cambium and then into the xylem. Lamellae, which restrict the movement of water and solutes through the apoplastic pathway are found in roots (Miyamoto et al., 2001) (Franke and Schreiber, 2007), (Ranathunge et al., 2003).

THE PATHWAY OF WATER

![Diagram of water transport pathway in the roots of a plant](http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/X/Xylem.html)

**Figure 1.1 Water transport pathway in the roots of a plant.**
Figure from Transport of Water and Minerals in Plants, 2010 (retrieved from [http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/X/Xylem.html](http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/X/Xylem.html) January 2010).

At any level, the water can leave the xylem and pass laterally to supply the needs of other tissues. At the leaves, the xylem passes into the petiole and then into the veins of the leaf. Water leaves the finest veins and enters the cells of the spongy and palisade layers. Here some of the water may be used in metabolism, but most is lost in transpiration.

This study will concentrate on plasma membrane intrinsic proteins (PIPs), which constitutes the largest plant aquaporin subfamily with 13 in Arabidopsis
(Johanson et al., 2001), 11 in Maize (Chaumont et al., 2000) and 24 in wheat (Forrest and Bhave, 2008).

### 1.2 Aquaporins

Water transport across the plasma membrane and the plant water relations have been studied for many years. The protein that transports water in biological cells is known as aquaporins (AQPs). Aquaporins are integral membrane proteins in the cell membrane from a larger family of major intrinsic proteins (MIP). They form pores in biological cell membrane and allow passive transport of water (Agre et al., 1993). The studies by (Weig et al., 1997, Kirch et al., 2000, Jang et al., 2004, Alexandersson et al., 2005) revealed that aquaporin expression could be up or down regulated by abiotic stresses such as cold, drought, salinity and abscisic acid (ABA), which suggests the important roles of aquaporins in the plant-water relation. The evidence of the role of AQP's in plant water relation was demonstrated by the heterologous expression of the *Arabidopsis* gene for PIP1b (AtPIP1b) in tobacco resulting in significant increase in the plant growth rate, transpiration rate, stomatal density and photosynthetic efficiency under favourable growth conditions (Aharon et al., 2003). In monocots the expression of many plasma membrane intrinsic protein (PIPs) from rice were up regulated in the upland rice during drought, whereas the corresponding genes remained unchanged or down regulated in the lowland rice under drought conditions (Lian et al., 2006, Lian et al., 2004). Therefore the regulation of AQPs in plants was related to drought tolerance mechanism. However the role of aquaporins in the plant-water relations still remains unclear. Though AQPs of Arabidopsis, maize, and rice have been extensively studied, limited information is known about the wheat aquaporins.
1.2.1 Historical Background of Aquaporin Research

Aquaporins are a class of proteins that mediate the passive movement of water and have been found in various cellular membranes (Maurel et al., 2002). The first aquaporin (CHIP28, renamed later AQP1) was discovered as a water channel from human erythrocyte membranes by Agre and his collaborators (Preston et al., 1992). Before the discovery of aquaporin it was believed that water permeates through the membrane by diffusion across the lipid bilayer, however several factors led to the prediction that water flows through membrane pores (Ray, 1960). The first aquaporin was isolated from the bovine lens cell and was originally called ‘major intrinsic protein’ (MIP), but later renamed AQP0 (Gorin et al., 1984), but its water transporting function was not determined until after the first human aquaporin was identified and expressed in frog oocytes (Preston et al., 1992). To date 13 aquaporins have been identified in mammals (AQP0 – AQP12). Availability of water is vital for all fundamental life processes of plants, e.g., seed germination, soil water uptake, photosynthesis, cell turgor, stomatal movement, and phloem loading, and MIPs appear to be involved in many of these processes, possibly by regulating the transcellular movement of water via symplast (diffusion of water through cytoplasm) (Johansson et al., 2000).
1.3 Aquaporin structure

1.3.1 Typical structure of an aquaporin monomer

Figure 1.2 The membrane topology of human AQP1.

AQPs are small and highly hydrophobic transmembrane proteins. Most of them are 250 – 300 amino acids in length with a molecular mass between 27 and 31 KDa. Each AQP has six transmembrane α helices (TMD1 – TMD6), which are connected by five loops (A-E), N- and C-terminal (Carboxyl terminal) domains protruding into the cytoplasm (Preston and Agre, 1991, Preston et al., 1994), refer figure 1.2. The amino acid sequence shows an internal symmetry in aquaporin proteins. The N-terminal (amino terminal) half is homologous to the C-terminal half of the protein, and both are inversely connected by the loops (Ralf et al., 2008). The N- and C-termini, are diverse in nature among aquaporin homologues (Santoni et al., 2006). The cytosolic loop B present in-between the second and third transmembrane domain and the extra-cytosolic loop E present in between the fifth and sixth
transmembrane domain possess a highly conserved NPA motif, composed of three amino acid residues, asparagine (N), proline (P), and alanine (A). The asparagine (N) is the key residue for the formation of the pore water selectivity filter. This NPA motif is conserved for all AQPs with the rare exceptions, as in some Arabidopsis and maize NIP (NOD26-like intrinsic proteins) and SIP isoforms, the alanine residue is replaced by leucine (L), valine (V), serine (S), threonine (T) or cysteine (cys/C) residues (Chaumont et al., 2001, Johanson et al., 2001). Many MIPs including plants and animals also contain a highly conserved AEF motif of unclear function, located within TMH1 (Zardoya and Villalba, 2001).

1.3.2 Aquaporin tetramers

Aquaporins usually function as tetramers in which each monomer forms an independent water channel. Structural analysis of two plant aquaporins, Kidney bean (Phaseolus vulgaris) PvTIP3;1 (α-TIP) and spinach (Spinacia oleracea) leaf SoPIP2;1, confirmed the typical tetrameric conformation of aquaporin in animal and bacteria (Daniel et al., 1999; Fotiadis et al., 2001; Karlsson et al., 2003). A tetrameric structure is essential for protein folding, stability and/or transport to the target membranes (Chaumont et al., 2005).

The ‘hourglass‘ model (Jung et al., 1994) (Fig.1.3) demonstrates the folding of TM1 – 3 and TM4-6 to form a symmetrical pore. The cytosolic loop B and extracellular loop E fold inward to meet in the center of the pore from opposite sides, bringing the NPA motifs into alignment and forming two half-helices. The two positively charged asparagines of the NPA motif form one wall of the pore when the prolines interact with each other. This is one of two major regions responsible for selectivity of the pore and gives the aquaporin its ion impermeability.
Figure 1.3 The hourglass model for the folding of human AQP1.

Each AQP1 monomer consists of six membrane-spanning α-helices with both N and C-termini located in intracellular. Water selectivity depends on the highly conserved NPA motifs and two pore helices. Adapted from (Spring et al., 2009).

1.4 Plant aquaporins and their molecular diversity

The diversity of plant aquaporins is higher than in other organisms due to higher degree of compartmentation of plant cells and a greater necessity for fine-tuned water control to adapt to environmental stress (Johanson et al., 2001). The first aquaporin identified in plants was γ-TIP (Tonoplast Intrinsic Protein) of Arabidopsis thaliana (Maurel et al., 1993). Subsequently, homologs of TIP were named PIPs, were found to be present in the PM (Plasma membrane) of A. thaliana (Daniels et al., 1994, Kammerloher et al., 1994). Plant aquaporins constitute a large and diverse family with 35 members identified in Arabidopsis thaliana (Johanson et al., 2001), 33
Plant aquaporins are divided into four subfamilies: Plasma-membrane intrinsic protein (PIP), tonoplast intrinsic protein (TIP), nodulin 26-like intrinsic protein (NIP), small basic intrinsic protein (SIP) (Johanson et al., 2001, Zardoya, 2005, Johanson, 2008, Johanson and Gustavsson, 2002) based on sequence similarities. GlpF-like intrinsic proteins (GIP), hybrid intrinsic proteins (HIP), and X intrinsic protein (XIP) are three new aquaporin subfamilies recently described in the moss Physcomitrella patens (Danielson and Johanson, 2008, Gupta and Sankararamakrishnan, 2009). The HIP subfamily was identified only in *P. patens*, whereas XIP appears to be in a variety of dicotyledonous plants (Danielson and Johanson, 2008).

Aquaporins function in various processes, including root water uptake, stress recovery, CO₂ transport, cell expansion and osmotic control (Maurel et al., 2002). Fluid transport across cellular barriers is of fundamental importance in animal and plant physiology. Physiological and agronomical attention has been focused on the mechanism of root water uptake and water transport to the shoot, as this determines cell growth and plant yield. Plasma membrane and tonoplasts (vacuolar membrane) in plant cells mainly depends on two types of plant aquaporins: plasma membrane intrinsic proteins (PIPs) and tonoplast intrinsic proteins (TIPs), respectively, although the membrane lipid compositions and other membrane proteins can also affect water permeability (Katsuhara and Hanba, 2008). Some plant aquaporins specifically transport water molecules, but others can mediate other low molecular weight compounds such as CO₂ (Flexas et al., 2006, Hanba et al., 2004, Uehlein et al., 2003), silicon (Ma et al., 2006), boron (Takano et al., 2006), ammonia (Jahn et al., 2004), and H₂O₂ (Bienert et al., 2008).
1.4.1 PIPs: Plasma Membrane Intrinsic Proteins

PIPs are Plasma membrane intrinsic proteins localized at the plasma membrane of the plant. The PIP subfamily can be further subdivided into two groups, namely, PIP1, with low water permeability, and PIP2 with high water permeability, as determined by functional testing using *Xenopus* oocytes (Chaumont et al., 2001, Katsuhara et al., 2002). Swelling assays in *Xenopus* oocyte showed that PIP2 aquaporins increase the water permeability of oocyte membranes 10- to 20-fold, whereas, PIP1 aquaporins do not affect the membrane water permeability considerably (Kammerloher et al., 1994, Biela et al., 1999). Heterotetramerization of PIP1 and PIP2 of plant aquaporins increase the water permeability of oocytes, supporting the notion of cooperative effects (Moshelion et al., 2002, Fetter et al., 2004). The transport of PIP1 proteins to the plasma membranes appears to be dependent on PIP2 expression (Zelazny et al., 2007). Generally, PIP1s have longer N-terminal and shorter C-terminal sequences compared to PIP2s, and these regions are divergent amongst PIP proteins (Chaumont et al., 2000).

1.4.2 TIPs: Characteristic vacuolar membrane aquaporins

TIPs are tonoplas intrinsic proteins, localized at the intracellular organelle (tonoplast). However, the cellular localization of some TIPs appeared to be complex and not restricted to vacuolar membrane (Barkla et al., 1999). Many TIPs exhibit specific water channel activity, while some transport both water and glycerol or ammonia or urea, although the *in vivo* roles of some are unclear. TIPs appear to control water exchange between cytosol and vacuole, their indispensability proven by RNA(i)-targeting of a TIP resulting in plant (*A.thaliana*) death (Ma et al., 2004). Functional analysis of yeast expressed *TaTIP2-2* by means of stopped-flow spectrometric studies provides clear evidence for water and NH$_3$ conductivity, rather than conductivity for NH$_4^+$. Results from inhibitor studies strongly suggest that NH$_3$ is not transported in file with water, but through a separate pathway, which could be supplied by the 5th central pore in a tetramer conformation (Bertl and Kaldenhoff, 2007).
1.4.3 SIP (Short basic intrinsic proteins)

The subcellular aquaporins in plants are called SIP (Short basic intrinsic proteins) because of their shorter sequences and abundant positively charged amino acids. The downstream cysteine (C) in the second NPA boxes, which is completely conserved in other subcellular aquaporins, are not found in the plant SIPS (Ishibashi, 2006). SIPs can be divided into two subgroups: SIP1 and SIP2.

The fourth subgroup of *Arabidopsis* aquaporins, small and basic intrinsic proteins (SIPs) were expressed in yeast and the water channel activity of membrane vesicles was measured with a stopped-flow spectrophotometry (Ishikawa *et al.*, 2005), revealing that SIP1.1 and SIP1.2, which are highly homologous with each other (70% identity) both transport water when expressed in yeast vesicles. On the other hand, SIP2.1 might act as an ER channel for other small molecules or ions. Immunoblotting using tissue subcellular fractionation showed their localization at the ER. All the SIP proteins were localized in the ER and not detected in the Plasma or Vacuolar membrane. GFP-tagged SIPs were also targeted to the ER, especially to the rough ER where protein synthesis occurs (Ishikawa *et al.*, 2005). As SIPs are not expressed in the plasma membrane, it is difficult to study their function.

1.4.4 NIP (NOD26-like intrinsic protein)

Soybean Gm NIP1-1 was the first identified and biochemically investigated plant MIPs (Fortin *et al.*, 1987). NIPs are highly diverse in nature with many different isoforms and different substrate specificities. The amino acid sequence of NIPs shows great diversity in the selective filters and are specifically involved in the transport of metalloids, such as silicic acid, arsenite, or boric acid in planta (Ludewig and Dynowski, 2009). Most of the NIPs are found in plasma membrane or endoplasmic reticulum. The plant NIPs/SIPs are most similar to the bacterial GlpF and the plant PIP and TIP are more similar to the animal AQPs (Quigley *et al.*, 2001). Based on the sequence similarity of the aromatic/arginine (ar/R) constriction region (Wallace and Roberts, 2004, Mitani *et al.*, 2008); the NIP subfamily is further
subdivided into three subgroups (NIP I, II, and III). These subgroups have different transport substrates, NIP I proteins transport water, glycerol (Dean et al., 1998), and lactic acid (Choi and Roberts, 2007), NIP II proteins are permeable to larger solutes, such as urea (Wallace and Roberts, 2005), formamide (Wallace and Roberts, 2005), and boric acid (Takano et al., 2006). The NIP III proteins are specific for silicic acid and its members transports silicon, although transport activities for other solutes such as undissociated arsenite and selenite have also been reported recently (Zhao et al., 2010, Ma et al., 2006, Ma et al., 2007, Ma et al., 2008).

1.5 Conserved residues of PIP aquaporin Family

The Valine (V) residue next to the first NPA motif is highly conserved in the aquaporin family and is considered to be essential for formation of a narrow water pore. Two NPA motifs together form a constriction that mediates a single-file movement of water molecules in either direction and are thought to be important for proton exclusion and the “global orientation tuning” of water molecules as they traverse the pore (Murata et al., 2000a) (Tajkhorshid et al., 2002). The first NPA motif is located in the LoopB and the second NPA motif is located in the LoopE (Fig 1.2). A second narrower constriction is formed by the aromatic/arginine (ar/R) selectivity filter, a tetrad residues, one each from TMH2 (residue H2) and TMH5 (residue H5). PIP2 aquaporins showed consistent increase in water permeability (Pf) when expressed in *Xenopus laevis* oocytes (Chaumont et al., 2000). In contrast most PIP1s showed low or no water transport activity. Phosphorylation of certain serine residues activates some PIPs and dephosphorylation rapidly closes them (Tornroth-Horsefield et al., 2006). This aquaporin inactivation can prevent water loss from cells under drought or strong salt stress (Kjellbom et al., 1999). This activation or inactivation can be effective for short-term (within a few hours) adjustment of the cellular water balance. The consensus phosphorylation site Ser274 in the C-terminal region is conserved in all PIP2 isoforms, independent of species, but not in PIP1 isoforms, which have a
shorter C-terminal region, whereas the consensus site at Ser115 in the first cytosolic loop is conserved in all PIPs, i.e in all PIP1 as well as PIP 2 isoforms (Johanson et al., 2001).

Interestingly, the substitution of isoleucine (I), Ile(244) of RsPIP1;3 with Val increased its water permeability to 250% of the wild type RsPIP1;3 (Suga and Maeshima, 2004). On the other hand, exchange of Val (235) of RsPIP2;2, which corresponds to RsPIP1;3 Ile (244), with isoleucine caused a marked inactivation to 45% of the original RsPIP2;2. This residue is located in loop E which dips into the membrane and forms part of the pore.

1.6 Spinach aquaporins

1.6.1 Water transport of spinach PIPs and gating of plant AQP

The gating of aquaporins is carried out to stop the flow of water through the pore of the protein. This may occur as a response to a number of stimuli, for example when the plant contains low amounts of cellular water due to drought. The gating of an aquaporin is carried out by an interaction between a gating mechanism and the aquaporin which causes a conformational change in the protein, blocks the pore and thus stops the flow of water through the pore. In plants it has been seen that there are at least two forms of aquaporin gating.

These are gating by the dephosphorylation of two highly conserved serine residues, which has been seen as a response to drought stress (Johansson et al., 1998, Johansson et al., 1996), and the protonation of specific histidine residues in response to flooding or a drop in the pH within the cell due to oxygen shortage during flooding (Tournaire-Roux et al., 2003). The phosphorylation of an aquaporin has also been linked to the opening and closing of plant aquaporin gate in response to temperature (Fig 1.4). The gating of aquaporin occurs by either Hg\(^{2+}\) or by low pH in some aquaporins (Yasui et al., 1999).
The spinach (Spinacia oleracea) aquaporin PM28A is the most abundant proteins of spinach leaf plasma membrane. The results obtained from the Xenopus oocyte and site directed mutagenesis suggested that phosphorylation at Ser274 in the C-terminal region regulates water channel activity of So PIP2;1 and another residue ser115 in the first cytosolic loop B is a consensus phosphorylation site, identified as an additional putative regulatory phosphorylation site for protein kinase A, which is conserved among putative plasma membrane MIP homologs and blocks the pore (Johansson et al., 1996).
Figure 1.4 Model of PM28A (spinach leaf plasma membrane aquaporin) and mechanism of plant AQP gating by blocking the water pore.

(A) PM28A, a major intrinsic protein showing potential phosphorylation sites on the cytoplasmic side of the membrane, six predicted transmembrane regions, the two highly conserved NPA boxes and Serine residues as potential phosphorylation sites are labelled with their amino acid numbers. Figure adapted from (Johansson et al., 1998).

(B) Phosphorylation of a specific residue (Serine-115) in the spinach AQP (SoPIP2;1) results in a change in conformation of one of the cytoplasmic loops (loop D, shown in 2 different conformations in red and blue, respectively). Loop D (blue) blocks the entrance of the water pore in the closed form. Phosphorylation of Serine-115 displaces the loop D from the pore (blue) into a new conformation (red), allowing water molecules to access the water pore. Figure adapted from (Wang and Tajkhorshid, 2007).
1.7 Rice aquaporins

1.7.1 Localization of Expression of ten OsPIP genes in rice seedlings

The transcripts of ten OsPIP gene in the leaves and roots of 2-week-old rice seedlings were examined by Guo et al., 2006, using quantitative RT-PCR. Table 1.1 denotes the expression of three genes, OsPIP1-3, OsPIP2-2 and OsPIP2-7, which were root specific, whereas OsPIP2-6 was found to be leaf specific and the expression of other six genes OsPIP1-1, OsPIP1-2, OsPIP2-1, OsPIP2-4, OsPIP2-5 and OsPIP2-3 were detected both in the leaves and roots (Guo et al., 2006b). However the expression of OsPIP1-1, OsPIP1-2 and OsPIP2-1 were higher in leaves, whereas OsPIP2-4 and OsPIP2-5 were higher in roots (Sakurai et al., 2005).

<table>
<thead>
<tr>
<th>S.No</th>
<th>Leaf Specific</th>
<th>Root specific</th>
<th>Expression in both leaves and Roots</th>
<th>Increased Expression in leaves at early seedling stage</th>
<th>Increased Expression in Roots at early seedling stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>OsPIP2-6</td>
<td>OsPIP1-3</td>
<td>OsPIP1-1</td>
<td>OsPIP1-1</td>
<td>OsPIP2-4</td>
</tr>
<tr>
<td>2.</td>
<td>OsPIP2-2</td>
<td>OsPIP1-2</td>
<td>OsPIP1-2</td>
<td>OsPIP1-2</td>
<td>OsPIP2-5</td>
</tr>
<tr>
<td>3.</td>
<td>OsPIP2-7</td>
<td>OsPIP2-1</td>
<td>OsPIP2-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td>OsPIP2-3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td>OsPIP2-4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td>OsPIP2-5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1 Organ specific expression rice aquaporin genes. shows the organ specific expression of 10 OsPIP genes, quantified by reverse transcription-PCR (RT-PCR) (Sakurai et al., 2005).
1.8 Wheat aquaporins

The PIP family has many members; for example there are 13 PIPs in Arabidopsis, 13 in Maize, and 11 in rice (Chaumont et al., 2001, Johanson et al., 2001, Sakurai et al., 2005). In wheat (Forrest and Bhave, 2008) have identified 24 MIPs, which includes twelve PIP1s and seven PIP2s (refer table 1.2). Two PIP1s were pseudogenes and the rest ten PIP1s were functional gene. (Yousif and Bhave, 2011) identified five genes designated TaPIP1-5b, TaPIP2-9b, TaPIP2-2, TaPIP2-2a and TaPIP2-2b, where TaPIP1-5b and TaPIP2-9b are homeologs of the previously reported TaPIP1-5 (EU177550) and TaPIP2-9 (EU177562). The database mining for the complete cds (coding sequence) of wheat PIP aquaporin resulted in ten complete sequences of TaPIPs from T.aestivum (accession numbers: AF139814 (TaPIP1), AF139815 (TaPIP2), AF139816 (TaPIP3), AF366564 (TaPIP1), AF366565 (TaPIP2); DQ867075 (TaPIP1), DQ867076 (TaPIP2), DQ867077 (TaPIP2) and DQ867078 (TaPIP2) from (Zhang et al., 2008)and GQ452384.1(TaPIP) and two from T.Turgidum (Ayadi et al., 2011) (accession numbers: EU162655 (TtPIP2) and EU162654 (TtPIP1) were found, which represents three subgroups of PIPs, PIP1, PIP2 and PIP3 respectively, based on the annotation.

1.8.1 Water transport of wheat PIPs

PIPs are sub classified into two subgroups PIP1 and PIP2, based on the difference in the deduced amino acid sequences, especially in the N- and C-terminals. Generally PIP2s have shorter N-terminal and longer C-terminal sequences compared to PIP1 (Forrest and Bhave, 2008). Water permeability of the tonoplast is higher than that of the plasma membrane as a result of the abundance of TIPs (Tyerman et al., 1999), and because water permeability of the whole system principally obeys the water permeability of the lower conducting part (that is, the plasma membrane), PIPs are the most important factors in cellular water uptake or water loss. This is the reason why PIPs have been the most intensively investigated in terms of the water relations of plant cells under various environmental conditions. The general feature of plant PIP
aquaporins is that PIP2 aquaporins show increase in water permeability (Pf) in the *Xenopus laevis* oocytes, when compared to PIP1 (Chaumont *et al.*, 2000). The known wheat AQPs are summarized in the table 1.2.

1.8.2 Regulation and Physiological roles of PIPs in *Triticum aestivum*

Aquaporins function in the regulation of cellular and subcellular pH homeostasis, maintenance of proton gradients, calcium signalling, bulk carbon transport and ROS or redox signaling (Murphy *et al.*, 2011). The regulation and function of AQPs are strongly associated with physiological responses to environmental signals and AQPs are regulated by multiple mechanisms, which includes gene transcription regulation, reversible posttranslational modifications, dynamic subcellular localization, protein protein interactions and hetero-tetramerization (Murphy *et al.*, 2011).
Table 1.2 Genes isolated from gDNA (genomic DNA) of wheat by (Forrest and Bhave, 2008) and the corresponding accession number.

PS – Partial sequence, ‘-’ sequence incomplete; therefore, presence/absence of intron is unknown.
1.9 Other functions of aquaporins

1.9.1 Aquaporins as CO$_2$ transporters

![Diagram of CO$_2$ transport]

**Figure 1.5** Transport of ambient CO$_2$ to the chloroplast through several barriers. 

1 leaf surface with stomata; 2 intercellular space; 3 cell wall; 4 plasma membrane of mesophyll cell; 5 cytoplasm; 6 chloroplast. An aquaporin can mediate CO$_2$ transport in the plasma membrane (*right path*) or via simple diffusion (*left path*). Internal conductance of CO$_2$ involves steps 2 to 6 mentioned above. Tissue and cellular structures are not to scale (Katsuhara and Hanba, 2008). Image adapted from (Katsuhara and Hanba, 2008).

The most important and basic function in plants is photosynthesis which can be limited by the supply of CO$_2$ to the chloroplast. The first evidence of transport of CO$_2$ in plants was reported in barley (Hordeum vulgare) (Figure 1.5) AQP, *HvPIP2;1*, transports CO$_2$ in addition to water (Katsuhara and Hanba, 2008). *HvPIP2;1* was introduced into the rice plants and the leaves of the transgenic rice plants showed a 40% increase in $g_i$ (internal conductance for CO$_2$ diffusion) compared to the leaves of wild-type rice plants (Hanba et al., 2004).

The same phenomenon was observed in tobacco plants: the CO$_2$ assimilation rate increased in leaves over-expressing *AtPIP1;2* (PIP1b) (Aharon et al., 2003) or *NtAQP1* (Flexas et al., 2006, Hanba et al., 2004,
Uehlein et al., 2003). These results suggest that the enhancement of aquaporin activity is a potentially promising way to promote plant CO$_2$ assimilation via improving g$i$ (Katsuhara and Hanba, 2008). Mammalian AQPs also could permeate CO$_2$ (Nakhoul et al., 1998, Cooper et al., 1998). The CO$_2$ assimilation rate enhancement in the transgenic tobacco may have been due to greater water conductance rather than CO$_2$ transport through aquaporin.

1.9.2 Aquaporins as H$_2$O$_2$ transporters

H$_2$O$_2$ is an important signaling molecule in the plant cells. Under environmental stresses, such as cold stress (Lee et al., 2004), H$_2$O$_2$ is generated on the outer surface of the cell, enters the cells and then causes intracellular reactions. H$_2$O$_2$ enters the cytoplasm through simple diffusion across the plasma membrane. Aquaporin mediated H$_2$O$_2$ diffusion has been monitored using human AQP8 and Arabidopsis TIP1;1 and TIP 1;2 by a fluorescence assay using an intracellular ROS sensitive fluorescence (Bienert et al., 2006, Bienert et al., 2007). The mean diameter of the pore size of the AQPs transporting H$_2$O$_2$ is 0.25 – 0.28 nm (Henzler and Steudle, 2000) and hAQP1 is 0.30 nm (Murata et al., 2000b), compatible with passage through the AQPs. The aquaglyceroporin, GlpF from E.coli has an wider pore size of 0.40 nm (Fu et al., 2007).

1.10 Heterologous expression systems

1.10.1 Introduction

As the function of proteins are difficult to study in their native tissues, studies on protein expression and structural biology strongly rely on heterologous expression systems for the production of recombinant proteins. Various heterologous expression systems are being used in recombinant protein production including bacteria, yeast, insect, and mammalian cells. Recently cell free translation systems have become available for general use. Today the most frequently used systems are based on expression in Escherichia coli and Baculovirus-infected insect cells. However, the situation
is different for membrane proteins. The insect cell system is superior to bacterial and yeast expression systems for protein processing and post translation modifications, particularly for membrane proteins (Yang et al., 1997). The eukaryotic host system seems better for producing functional, membrane-inserted eukaryotic proteins than prokaryotic hosts. This could reflect a slower rate of translation and folding in eukaryotic cells compared with bacteria. To date, the recombinant production of membrane proteins is still a matter of trial and error (Grisshammer, 2006).

1.10.2 Bacterial expression system

1.10.2.1 Expression in E.coli

*Escherichia coli* has been extensively used as a heterologous expression system for protein production (Daly and Hearn, 2005). The factors that make *E.coli* a system of choice for expression studies are, it is a simple, fast, cheap and safe expression system for large-scale production of different types of proteins (Lundstrom, 2007). Although a number of high-resolution structures have been solved by recombinant expression of soluble proteins in the prokaryote *E.coli*, membrane proteins have been less successful. The main problem with this system is the lack of intracellular machinery to achieve post-translational modification(s) of the expressed eukaryotic proteins to fold correctly and to obtain a functional protein. The protein product may be obtained as insoluble, miss-folded inclusion bodies, so that subsequent solubilization and re-folding steps are required (Makrides, 1996, Marston, 1986). As a result of inadequate intracellular chaperone concentration, incorrect folding can occur (Cole, 1996) or the reducing environment of the cytoplasm (Bardwell, 1994, Li et al., 2001, White et al., 1994) may result in formation of inappropriate disulphide links during purification. Many membrane proteins seem to be highly toxic to the bacterial host cells because of their transmembrane topology and therefore bacterial growth is restricted thereby reducing the yield of recombinant proteins. The proteins with high level of disulphide bridges (White et al., 1994) or proteins that require posttranslational modifications such as glycosylation, proline
cis/trans isomerization, disulphide isomerization, lipiddation, sulphonation or phosphorylation are generally not suitable to express in *E. coli* (Lueking *et al.*, 2000). The proteins expressed in *E. coli* retain their amino-terminal methionine, which may affect the stability (Chaudhuri *et al.*, 1999, Takano *et al.*, 1999) and cause immunogenicity. The function of certain proteins can be altered because of the inability of the *E. coli* to glycosylate recombinant proteins (Meldgaard and Svendsen, 1994, Jenkins *et al.*, 1996). An example of this limitation is the expression of erythropoietin (EPO) in *E. coli*. EPO is a glycoprotein that when expressed in *E. coli* in the unglycosylated form is less resistant to unfolding than its native glycosylated counterpart (Narhi *et al.*, 1991). Refolding of a recombinant protein present as inclusion bodies is often required for *E. coli* expressed proteins (Majerle *et al.*, 1999, Chen and Hai, 1994, Koganesawa *et al.*, 2001). The optimization of folding conditions has to be done carefully. Developing such folding conditions may be difficult to achieve and time consuming (Wang *et al.*, 2000, Tsujikawa *et al.*, 1996).

Because of the limitations, *E. coli* has not become an expression system of choice for plant membrane proteins. However a number of bacterial membrane proteins have been expressed successfully in *E. coli* and subjected to purification and structurally analyzed (Lundstrom, 2006). Successful expression of eukaryotic membrane protein requires modifications such as deletion and mutation of the gene sequence and application of appropriate fusion partners and purification tags. Other bacterial expression systems such as *Lactococcus lactis*, *Bacillus subtilis* and *Halinobacterium salinarum* have also been used for the production of recombinant proteins in addition to *E. coli* (Lundstrom, 2007), but *L. lactis* resulted in the less production of total membrane protein, *B. subtilis* has high endogenous protease activity and also restricted to secreted proteins has limited the use of this system and it was essential to use Bop (Bacterio-opsin protein) as a fusion partner with *H. salinarum* which resulted in less protein production.
1.10.3 Expression in yeast cells

Yeast expression systems have many advantages over *E.coli* expression system, by their easy use, large-scale production of recombinant proteins and their eukaryotic post translation machinery. Two yeast strains used for recombinant protein expression are *Saccharomyces cerevisiae* and *Pichia pastoris* (methylotrophic yeast). Proteins that cannot be expressed in *E. coli* with the correct level of post-translational modification to obtain functionally active protein have been subsequently expressed in the methylotrophic yeast *Pichia pastoris* (Monsalve et al., 1999).

Spinach aquaporin *Spinacia oleracea* SoPIP2;1 was the first structure solved of heterologously expressed aquaporin in any eukaryotic species. SoPIP2 was expressed both as His-tagged and as non-tagged protein in the methylotrophic yeast. Both the tagged and untagged clones express SoPIP2;1 variants at similarly high levels. Recombinant SoPIP2;1 showed efficient water channel activity, when both the tagged and untagged proteins were reconstituted into proteoliposomes and exposed to an osmotic gradient (Kukulski et al., 2005). Several crystal forms were obtained after reconstituting the His-tagged and non-tagged SoPIP2;1. The fourth subgroup of *Arabidopsis* aquaporins, the small and basic intrinsic proteins (SIPs) have also been expressed in yeast (Ishikawa et al., 2005).

1.10.4 Expression in insect cells

1.10.4.1 Baculovirus expression system

Recombinant baculovirus infection of insect cell lines is the second most commonly used expression system after *E. coli* (Luque and O’Reilly, 1999). Insect Baculoviruses are used for the heterologous expression of eukaryotic proteins in Sf9 or Sf6 cells under the control of the strong viral polyhedrin promoter. The baculovirus system has been used extensively to express soluble proteins, and more recently to express functional membrane transporters and channels, including the cardiac Na+-Ca2+ exchanger Na+/glucose cotransporter (Smith et al., 1992), CFTR Cl-channel (Oriordan et al., 1995) and renal Na+/phosphate cotransporter (Fucentese et al., 1995). The
insect cell system is in general superior to bacterial and yeast expression systems for protein processing and posttranslational modifications, particularly for membrane proteins. However, membrane protein stability and yields from insect cells are often substantially lower than yields for soluble proteins, possibly because of inefficient recognition of heterologous signal peptides by the insect cell protein translocation machinery cell toxicity and/or the limited amount of membrane available for protein insertion (Yang et al., 1997). Baculovirus infected insect cells of the moth *Spodoptera frugiperda* (Sf9 and Sf21 cell lines) have been used for biochemical and electrophysiological characterization of plant K$^+$ channels (Gaymard et al., 1996). Mammalian aquaporins expressed and purified in Sf9 cells include AQP2, AQP4 and AQP11 (Werten et al., 2001, Yang et al., 1997, Yakata et al., 2011). Therefore the insect cell expression system is suitable for functional expression and purification of other members of the aquaporin family.

### 1.10.5 Problems in heterologous expression system

In comparing the extensively used *Xenopus oocytes* with insect cells (Sf9 or Sf21), the insect cell are more amenable to the patch-clamp studies, because of the absence of the vitelline membrane around the cells. The inconveniences of the insect cells are the construction of expression vectors (recombinant baculovirus encoding for the gene of interest) and routine purification are difficult in comparison to other vectors and these expression vectors are limited to the Sf9 or Sf21 cells, whereas the vectors that function in *Xenopus oocytes* can also be used in cell types such as COS and CHO cells (Chinese hamster cells). Furthermore the Sf9 cells die within 1-3 d after infection and therefore allowing fewer experimental possibilities, while the oocytes can be used up to 1 week after injection of cRNA (complementary RNA) or DNA (deoxyribonucleic acid) vectors (Dreyer et al., 1999).
1.11 Functional testing of aquaporins expressed in heterologous expression systems

To assay the function of the aquaporins, stopped flow spectrophotometer has been used to measure the water permeability of the heterologously expressed aquaporins either in isolated vesicle (insect cells) fractions (Yang et al., 1997, Yakata et al., 2011, Werten et al., 2001) or in the yeast spheroplast (Laizé et al., 1999, Beitz et al., 2006, Daniels et al., 2006, Kojima et al., 2006, Sakurai et al., 2005, Suga and Maeshima, 2004). A stopped flow instrument is most frequently used rapid kinetics techniques in solution over timescales of about 1ms up to 100’s seconds. A stopped flow instrument is a rapid mixing device, where two or more solutions containing the reagents are mixed and the chemical kinetics of the reaction is studied. The freshly mixed reagents are stopped in an observation cell and the sample is irradiated with the monochromatic light and the reaction proceeds the change in the signal is recorded or the absorbance at a specific wavelength, is recorded as a function of time. The reaction rate can be determined by analyzing the kinetics of the reaction. (http://www.hitechsci.com/techniques/stoppedflow and http://www.photophysics.com/tutorials/stopped-flow-spectroscopy).

1.11.1 Functional testing of aquaporins in Yeast

1.11.1.1 Review of Light scattering

Light Scattering is a method based on the dependence of elastically scattered light (Rayleigh scattering) on cell volume. The osmotic water permeability is measured by light scattering, a well-established method in osmotically responsive liposomes, sealed membrane vesicles and some small cells. The water permeability was first measured in human red blood cell using Light scattering (Mlekoday et al., 1983) and subsequently in membrane vesicles, liposomes reconstituted with water channels, and cell suspensions (Verkman et al., 1985, Verkman, 1992, Dobbs et al., 1998). The presence of aquaporin proteins in the small vesicles allow a faster water uptake (Nishihara
The light scattering method is simple to apply and very small samples are required; however, there are potential problems in quantitative data interpretation, including cell/vesicle size heterogeneity, motion artifacts just after mixing which occur before flow stops, refractive index changes producing anomalous scattering signals, and in cells, unstirred layers, and scattering from intracellular structures (Verkman and Mitra, 2000).

1.11.1.2 **Functional testing of aquaporins in Yeast using Stopped-flow spectrometry**

Need to preface with some description of the need and method of preparation of protoplasts. Volume changes in yeast protoplasts, resulting from transmembrane water transport, were followed by 90° light scattering. This resulted from an outwardly directed osmotic gradient inducing water uptake and volume increase, which gives rise to a decrease in the intensity of the scattered light. Single-and multi-exponential functions were fitted to the data and the osmotic permeability coefficient $P_f$ was calculated using the following equation:

$$P_f = \frac{1}{\tau} \cdot \frac{V_0}{S_0} \cdot \frac{V_w}{C_{out}}$$

Where $\tau$ is the time constant for the exponential decay, $V_0$ the initial mean protoplast volume, $V_w$ the partial molar volume of water (18 cm$^3$ mol$^{-1}$), $S_0$ the initial mean protoplast surface area and $C_{out}$ the external osmolarity after the mixing procedure. The initial diameter of the protoplasts in 1.8M sorbitol was about 4.5 μm, as determined by light microscopy (Bertl and Kaldenhoff, 2007).

1.11.1.3 **Ammonia transport in the yeast cell**

Intracellular pH changes in intact yeast cells, resulting from uptake of NH$_3$ or NH$_4^+$, leads to changes in fluorescence intensity of fluorescein, which was monitored in a stopped-flow spectrophotometer (SFM 300, BioLogic) using 490 nm excitation in combination with a cutoff filter of 515 nm, resulting in an inwardly directed gradient for NH$_4^+$ and NH$_3$. (Bertl and Kaldenhoff, 2007), where the cells were equilibrated in the incubation buffer.
(50mM NaCl, 10 mM Tris/HCL, pH 8) and were mixed with an equal volume of the test solution containing (50 mM NH₄Cl, 1 mM Tris/HCL, pH 8).

1.11.2 Functional testing of aquaporins in *Xenopus laevis*

The water channel activity of AQPs was determined by using immature eggs (oocytes) of *Xenopus laevis*. Oocytes of *Xenopus* are impermeable to water, thus making them a possible expression host to test the water transport properties of AQPs from various other biological entities (Kataoka *et al.*, 2009, Duchesne *et al.*, 2003, Kikawada *et al.*, 2008, Philip, 2011). The corresponding mRNA (messenger RNA) is injected into the oocytes and translated, and then the rate of swelling or shrinking of oocytes is monitored. Though this method is widely used for determining the water channel activity of individual proteins expressed heterologously in *Xenopus* oocytes, there are some disadvantages. The expression levels vary with the oocyte and it is difficult to quantify the protein accumulated in each oocyte because of its small quantity. This method is applicable for the plasma membrane aquaporins but not for those targeted to endomembrane (Suga and Maeshima, 2004). Figure 1.6 represents the swelling of oocytes with AQP and oocytes without AQP remained the same over the time (0.5 min – 3.5 min).

![Figure 1.6 Agre's experiment with Xenopus oocyte cells containing or lacking aquaporin. Aquaporin is necessary for making the 'cell' absorb water and swell. Picture from (Agre, 2003)]
1.11.2.1 Water permeability in *Xenopus oocytes*

An important system for measurement of water permeability is the *Xenopus* oocyte (0.12 cm diameter), which is the principal system for heterologous water channel expression. The original swelling assay involved dilution of the extracellular solution with distilled water, and estimation of oocyte volume by measurement of two orthogonal oocyte diameters on a video monitor every 1–5 min (Fischbarg *et al.*, 1990). An improved quantitative imaging approach was subsequently developed in which the shadow cast by an oocyte was recorded and digitized using transmission light microscopy (Zhang and Verkman, 1991). Relative oocyte volume was computed from cross-sectional area by image masking and pixel summation. This approach has been utilized extensively to study new water channel cDNAs (Copy DNA), to test whether various transporting proteins contain aqueous pores, and to quantify single channel water permeabilities of aquaporins (Yang *et al.*, 1997). A recent advance in the design of an oocyte perfusion chamber permitted analysis of early oocyte volume changes in response to solute gradients (Meinild *et al.*, 1998).

1.12 Recombinant expression of membrane proteins

1.12.1 Factors necessary for the production of a functional eukaryotic membrane protein

The use of eukaryotic expression system (with slower rate of protein translation and folding (Widmann and Christen, 2000) would seem better for producing functional, membrane-inserted eukaryotic proteins than prokaryotic expression system (Grisshammer and Tate, 1995, Sarramegna *et al.*, 2003, Tate, 2001). An advantage of using eukaryotic expression system is recombinant proteins will undergo faithful post translational modification. But is there any general cloning strategy that guarantees the production of functional eukaryotic membrane protein? The answer is no. To date, the production of recombinant membrane proteins is still a matter of trial and error (Grisshammer, 2006). The important characteristics of membrane proteins to
consider for recombinant expression are *in vivo* turnover (mRNA and protein stability) or dependence on molecular chaperones specific for the proper folding of membrane folding (Grisshammer, 2006).

### 1.12.2 Multistep process in expressing a functional membrane protein

A eukaryotic protein in the plasma membrane must first be targeted as a nascent polypeptide to a protein-conducting channel (translocon) in the endoplasmic reticulum (ER) membrane, where topology of the membrane protein is decided. Then the polypeptide is released from the translocon laterally into the ER membrane, where membrane protein folds into its native conformation. Proteins that undergo correct folding assured by a quality control system pass further into the Golgi apparatus (Ellgaard *et al.*, 1999), although post-ER quality controls might also apply (Coughlan *et al.*, 2004). Proteins that don’t undergo folding will be degraded (Meusser *et al.*, 2005, Tsai *et al.*, 2002). Most of the plasma membrane protein trafficking occurs via the endoplasmic reticulum and golgi complex with glycosylation and modifications to the glycan chains being made en route to the cell surface. Finally, the membrane protein will reach the target plasma membrane (Grisshammer, 2006).

### 1.13 Regulation of aquaporin

#### 1.13.1 Regulation of plant aquaporin by phosphorylation and membrane trafficking

Aquaporin trafficking is regulated by phosphorylation of aquaporin protein, as observed in mammalian AQP2 (Brown, 2003). Plant aquaporins are targeted to certain membranes. Plant aquaporins are localized either in the plasma membrane, the vacuolar membrane (tonoplast) or in the peribacteroid membrane of nitrogen-fixing symbiotic root nodules found by immunocytochemistry, immunodetection and expression of green fluorescent protein (GFP)-fusion proteins (Fortin *et al.*, 1987, Morrison *et al.*, 1988, Daniels *et al.*, 1994, Kammerloher *et al.*, 1994, Robinson *et al.*, 1996,
FleuratLessard et al., 1997, Chaumont et al., 1998, Barkla et al., 1999, Barrieu et al., 1999, Cutler et al., 2000, Kirch et al., 2000, Reisen et al., 2003). The cellular distribution of some aquaporins is more complex than plasma membrane or tonoplast localization. To reach their final destination plasma membrane aquaporins are transported to the cell surface along the secretory pathway, traversing the endoplasmic reticulum (ER), Golgi apparatus and then into different types of vesicles depending on their membrane target(s) (Chaumont et al., 2005). Export of membrane protein from ER is a highly regulated process requires specific interaction with components of coat protein complex 2 vesicles (Sato and Nakano, 2007). Recent studies show that the diacidic motifs (D/E)x(D/E), with x representing any amino acid residue, which are usually found in the N- and C-termini of the proteins play an important role in the export of transmembrane proteins from the ER to the Golgi (Mancias and Goldberg, 2007, Nishimura and Balch, 1997, Otte and Barlowe, 2002, Votsmeier and Gallwitz, 2001, Marylou et al., 2007, Zuzarte et al., 2007).

1.13.2 Glycosylation of Aquaporins

Glycosylation is important for some membrane proteins to reach their target site. Glycosylation of human aquaporin AQP2 is essential for exit from the Golgi complex and sorting to the plasma membrane (Hendriks et al., 2004). The glycosylation of plant aquaporins have been little studied, however, it was observed that GmNOD26 is sensitive to endoglycosidase-H treatment (Miao et al., 1992) and in the ice plant TIP2-1 (Vera-Estrella et al., 2004). (Hove and Bhave, 2011), identified that NIPs and At TIP1-4 have a putative O-glycosylation site mostly in the loop B and O-glycosylation, G-H-I-S-G-A/G-H-X, unique to NIPs only, was also identified. The S132 residue was predicted as a glycosylation site only in NIPs and AtTIP4;1, although it was conserved in PIPs and TIPs. Hove and Bhave, 2011, also suggests that their prediction showed N-glycosylation sites in the N and C termini, loops A, B and E, or TM2, for some MIPs (e.g., AtPIP2;1, OsNIP2;1).
1.14 Aquaporin methylation

Methylation of proteins can occur as either N-methylation or carboxymethylation of amino acid residues such as lysine. A plant aquaporin methylation study in the *Arabidopsis* root plasma membrane revealed that the N-terminal tail of PIP exhibits multiple modifications. Mass Spectroscopy study showed that the initial methionine (M) is N-acetylated in PIP1s, whereas in PIP2s the initial methionine is cleaved. Therefore during post-translational modification the initial methionine are processed in all PIP isoforms (Santoni et al., 2006).

1.15 Protein purification

To characterize the structure of the protein, it generally requires access to highly homogenous and pure protein preparations. Various affinity tags have been introduced at the N- or C-terminal or even within the coding sequence of the gene of interest to facilitate the purification. The most commonly used tag is multi-histidine (either hexa or deca histidine), which binds to Ni$^{2+}$ and therefore allows purification based on immobilized metal affinity chromatography (IMAC). Other common purification tags are streptavidin (Strep), biotin, FLAG and hemagglutinin tags. Obviously when antibodies are available, affinity chromatography can be applied. Other means of purification include ammonium sulphate precipitation and sucrose gradients, although these methods require large quantities of material and might therefore not be suitable for recombinant proteins expressed at low levels or membrane proteins with low recovery yields. In addition, gel filtration, size exclusion chromatography, hydrophobic interaction and reverse-flow chromatography are methods to be considered for protein purification (Lundstrom, 2007).

Intrinsic membrane proteins require special conditions for purification. Because of their transmembrane topology, separation of proteins and lipids is necessary by the addition of detergents. Solubilization by detergents is a complicated process, and a vast number of detergents have been tested. In general, detergents are highly target-specific, which means that each target has
to be screened for appropriate detergents. Commonly used detergents are CHAPS (3-[(3-cholamidopropyl) dimethylaminio]-1-propane-sulfonate, Triton X-100, n-Octylglucoside, n-Nonylglucoside, n-Dodecylmaltoside and FOS-Choline and cocktails thereof.

1.16 Aims of this project

The major aim of the project is to clone and characterize the full length coding sequence of PIP subfamily proteins expressed from wheat aquaporin genes, and to test the function of PIP proteins using a heterologous expression system. The specific aims are

- To amplify, clone, sequence and analyze full length PIP genes from *Triticum aestivum* cv. Cranbrook from both genomic and transcript (cDNA) level.
- To further characterise the wheat PIP genes and their putative proteins by comparison with those in rice, *Arabidopsis* and identify their evolutionary relationship.
- To verify the tissue (leaf and root) and stage specific expression of cloned wheat aquaporin genes (PIPs) by analysing mRNA for the same gene in different tissues at different stages.
- To test the function of particular Aquaporin proteins from Wheat using heterologous expression system which requires: Subcloning the selected wheat Aquaporin gene and control human aquaporin 3 gene into baculovirus transfer vectors pBacPAK 9 (untagged) and pBacPAK His1 (His tagged) vector. To develop a novel method to test the activity of the wheat aquaporin in comparison to hAQP3 by measuring the whole insect cell volume changes using ImageJ software.
- Immunoblot analysis to verify the expression of the histidine tagged human AQP3 and wheat aquaporins TaPIP1-5 and TaPIP2-8.
- Extraction and Purification of membrane protein.
The attainment of the above aims will help to understand the function of aquaporin proteins in wheat and provide a method useful to express a plant aquaporin protein in heterologous system.

1.17 Thesis structure

Chapter 1
Introduction of aquaporin structure, function and expression systems used to study the function of the membrane protein.

Chapter 2
The first result chapter will describe the Identification and analysis of two full length TaPIP genes from wheat root cDNA comparisons.

Chapter 3
The second result chapter will describe the expression of wheat PIP genes in the Baculovirus infected insect cells.

Chapter 4
The third result chapter will describe the new method developed to test the function of the aquaporin.

Chapter 5
General conclusion and future directions.

Chapter 6
References

Appendix: Additional supporting data is available in the appendix.
Result Chapters
Cloning and sequencing of two TaPIP genes at both genomic and cDNA level
Chapter 2  Cloning and sequencing of wheat PIP genes

2  Cloning and sequencing of two *TaPIP* genes at both genomic and cDNA level

2.1  Introduction

2.1.1  PIPs

Phylogenetic analysis has revealed the separation of the PIPs into two subgroups, PIP1 and PIP2 (Kjellbom *et al.*, 1999). A further subgroup, PIP3 (Pp PIP1-3), is found in the *Physcomitrella patens* genome (Danielson and Johanson, 2008). PIP1 proteins have been reported to have low water permeability (and show glycerol and urea permeability), and PIP2 proteins transport water almost exclusively (Chaumont *et al.*, 2000, Moshelion *et al.*, 2002, Fetter *et al.*, 2004, Suga and Maeshima, 2004, Sakurai *et al.*, 2005). However, PIP1 proteins form heterotetramers with some PIP2 monomers to be able to facilitate water permeation (Fetter *et al.*, 2004, Zelazny *et al.*, 2007). The water permeability of the PIP1 proteins was increased by site directed mutagenesis (Suga and Maeshima, 2004), suggesting that residues in the helix loop E and the cytoplasmic C-terminus tails are involved in the water transport and the functional regulation differences.

2.1.2  Criteria for gene selection

The identification and analysis of partial *PIP* sequences in wheat were reported in (Forrest and Bhave, 2008). However, the full length coding sequence of wheat PIPs had yet to be identified as was necessary for cloning of the genes to study the function of the protein. Therefore the partial sequences of PIP previously identified by Forrest and Bhave, including nineteen wheat *PIP* genes (Genbank accessions EU177547-EU177565) and five wheat *TIP* genes (Genbank accessions EU177566-EU177570) (listed in table 1.2, chapter 1), were used as the starting point for searching databases for sequences for primer design to amplify full length PIP genes from cDNA.
Chapter 2  
Cloning and sequencing of wheat PIP genes

My study mainly focused on an attempt to develop new methods for assessing the transport activity of aquaporins. To enable the testing of new method two wheat AQPs were selected, one expected to have high water transport activity and the other low activity from comparison with the measured activities of their counterparts in rice or Arabidopsis. According to the literature, PIP1 aquaporins transport very little water compared to PIP2 aquaporins (Chaumont et al., 2000, Moshelion et al., 2002, Fetter et al., 2004, Suga and Maeshima, 2004, Sakurai et al., 2005). Therefore all the existing wheat AQP (PIP1s and PIP2s) in the database (EU177547 to EU177570) was compared with the rice microarray database (http://www.ricechip.org/). The PIP2 gene identified as PIP2:8 (EU177561) was chosen as the candidate high transport example. The primary amino acid sequence comparison revealed that the orthologous gene 8 in rice is OsPIP2-6 and the expression level of OsPIP2-6 is high in the root tissues of early seedling stage plant. TaPIP1-5 (EU177550) was chosen as the predicted low transport comparison for a study of water transport activity. The expression of TaPIP1-5 was analysed by Bhave and Forrest using quantitative or Real Time PCR. The study showed that TaPIP1-5 is differentially expressed under salt stress at different time points. The primary amino acid sequence comparison revealed that the TaPIP1-5 is orthologous to OsPIP1-1 and is expressed almost equally in both roots and leaf blades. The specificity for water transport was predicted by the analysis of the primary amino acid sequence shown in table 2.1 published by (Quigley et al., 2001). The glycerol transporting aquaporin carries tyrosine (Y) or phenylalanine (F) in the position 1 (P1) after helix 3 (H3) whereas the water transporting AQPs carries Threonine (T), Glutamine (Q), Methionine (M), Glutamic acid (E), Valine (V) and Alanine (A); Serine (S) and Alanine (A) in AQPs at P2 in loop B (LB) was replaced by Aspartic acid (D) in GLPs; Alanine (A) in AQPs at P3 after helix 5 (H5) was replaced by Arginine (R) and Lysine (K); Tryptophan (W) in AQPs was replaced by Isoleucine (I), Valine (V), Leucine (L), Methionine (M), Alanine (A) and Tryptophan (W); two NPA motifs present at LB and LE, shown in table 2.1. Thus this study aimed to amplify full length TaPIP1-5 and TaPIP2-8 genes from genomic
Cloning and sequencing of wheat PIP genes

DNA of *Triticum aestivum* (cv. Cranbrook) and from the tissues, root and shoot cDNA of 8 day-old early seedling stage and test the prediction of their transport properties.

<table>
<thead>
<tr>
<th>Consensus position</th>
<th>P1 after H3</th>
<th>P2 at NPA-LB</th>
<th>P3 after H5</th>
<th>P4 at NPA-LE</th>
<th>P5 at NPA-LE</th>
<th>NPA-LB</th>
<th>NPA-LE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQP-type</td>
<td>T/Q/M/E/V/A</td>
<td>S/A</td>
<td>A</td>
<td>Y/F</td>
<td>W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLP-type</td>
<td>Y/F</td>
<td>D</td>
<td>R/K</td>
<td>P/A</td>
<td>I/V/L/M/A/W</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.1 Consensus residues distinguishing AQPs and GLP proteins.**
The table adapted from (Quigley *et al.*, 2001). P1, P2, P3, P4 and P5 are positions 1, 2, 3, 4 and 5; LB and LE denotes Loop B and Loop E.

### 2.2 Materials and Methods

#### 2.2.1 Prepared solutions

##### 2.2.1.1 Buffers and solutions

All buffers were prepared using sterile dH₂O (deionised water). Unless otherwise stated, all buffers and solutions were prepared according to Sambrook and Russell (2001).

The solutions required for running a crystal violet gel were prepared according to the protocols provided in the TOPO® XL PCR Cloning kit (Invitrogen, Version M, 14 June 2004, 25-0199). All buffers used for RNA (ribonucleic acid) work were prepared with DEPC-treated dH₂O (RNase-free) or Milli-Q dH₂O (purified with a Milli-Q system to 18.2 MΩ cm at 25°C).

##### 2.2.1.2 Sterilisation

Sterilised glassware, plasticware and solutions were used. All glassware and most disposable plasticware were autoclaved (121°C for 20 mins). Some disposable plasticware was bought sterile, including 10ml, 15ml and 50ml Falcon tube (BD), barrier tips (Continental Lab Products, California, USA) and 25cm² T flask (BD Falcon).
2.2.2 Plant material

The wheat sample *T. aestivum* (cv. Cranbrook), a drought-sensitive variety (Ji *et al.*, 2011) in the table 2.2 was kindly provided as seeds by Dr. Greg Grimes at the Australian Winter Cereals Collection (AWCC; Tamworth, NSW, Australia).

<table>
<thead>
<tr>
<th>AUS Accession</th>
<th>Species</th>
<th>Cultivar</th>
<th>Genome</th>
<th>Ploidy and chromosome no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>22660</td>
<td><em>T. aestivum</em></td>
<td>Cranbrook</td>
<td>AABBDD</td>
<td>2n = 6x = 42</td>
</tr>
</tbody>
</table>

Table 2.2 Wheat cultivar used in the study

2.2.3 Sterilization of seeds

Seed sterilization was performed in a sterile tissue culture hood. Seeds were counted and placed in a 1.5 ml microcentrifuge tubes and 1ml of 15% bleach (in sterile ddH₂O) was added. Tubes were vigorously shaken or vortexed for 30sec to avoid seeds clumping together and then incubated for 20min with occasional rotation. Seeds were allowed to settle down and then the bleach solution was discarded and 1ml sterile ddH₂O was added to rinse out bleach. Again the seeds were allowed to settle and then ddH₂O was poured off, 1ml of 70% ethanol was added to the tubes and shaken continuously for 30-45sec. The seeds were allowed to settle at the bottom and ethanol was poured off. 1ml of sterile ddH₂O was added to the seeds, seeds were allowed to settle down for a while and sterile ddH₂O was discarded. The seeds were repeatedly washed with sterile ddH₂O for four times (Chatfield and Raizada, 2005).

Sterilized seeds were spread on the soil (Potting mix ratio, 5soil:4vermiculate:2.5perlite) in the pots and pots were covered with a plastic wrap. Additional dark cover was used to protect the plant from light exposure and incubated at 4°C for 3days for vernalization. After 3days the pots were placed under light in a Thermoline growth cabinet (16 hour day and 8 hour nights, 24°C and 80% humidity) for 5days (total
Chapter 2  Cloning and sequencing of wheat PIP genes

8 days from the incubation time). Leaves and roots were harvested after 8 days, snap-frozen in liquid nitrogen and stored at -80°C.

2.2.4 Plants growth condition

Wheat (cv. cranbrook) seeds were germinated for 3 d at 25°C, planted into pots filled with soil (Potting mix ratio, 5:1:1, soil:vermiculate:perlite) (Chatfield and Raizada, 2005) and grown in a Thermoline growth chamber under 16 h light/8 h dark and at day/night temperatures of 25/20°C at a relative humidity of 80%. Plants were grown in tap water for the first 14 d, and then were supplied with half strength MS salt mixture solution. Leaves and roots were harvested using liquid Nitrogen from 21-day-old plants (early tillering stage) and 8-day-old plants (early seedling stage). Shoots and roots were harvested after 21-day-old plants and 8-day-old plants, snap-frozen in liquid nitrogen and stored at -80°C. The harvested leaves and root materials were stored at -80°C until RNA and DNA extraction.

2.2.5 Bioinformatics Methods

The main criterion for the study was to investigate the water transport activity of the wheat MIP genes from *Triticum aestivum* (cv. Cranbrook). All the 24 wheat MIP genes were retrieved from the NCBI database using the accession numbers (Refer chapter 1, table 1.2) and used for further searches as described in the below subsections (2.2.5.1 and 2.2.5.2). The AQP protein sequences of Barley, Wheat, Rice and Arabidopsis were obtained from NCBI protein database to analyse the orthology between the genomes of different species.

2.2.5.1 *In silico* search methods to identify Transcript Assemblies (TAs) for PIP1 and PIP2

Different methods were employed to predict the wheat genes from ESTs (Expressed sequence tags), their properties, tissues of
expression and possible functions; these are summarised below. Below is the search using BLAST (Basic Local Alignment Search Tool) tool.

**blastn, tblastX and blastp using partial predicted wheat sequences**

The partial DNA sequences coding for wheat aquaporins (24 MIPs) were obtained from the NCBI database as reported in the study (Forrest and Bhave, 2008). A blastn search was performed on each partial wheat PIP cds (Gene bank accessions EU177547 to EU177570) within NCBI restricted to Tritium aestivum ESTs and within TIGR database (http://blast.jcvi.org/euk-blast/plantta_blast.cgi accessed June-July 2009), resulting in a list of wheat ESTs and Transcript assemblies (TAs) (Table 2.6). Only those TAs comprising at least ten or more ESTs were considered, and all the ESTs in the TA was analyzed for the full length coding sequence to produce functional aquaporins by analyzing the Open Reading frames (http://www.ncbi.nlm.nih.gov/projects/gorf/) for the deduced amino acid sequence. The EST with highest match was further used to perform a tblastx search against the rice database in TIGR database (http://blast.jcvi.org/euk-blast/plantta_blast.cgi); accessed June-July 2009, to find the homologous genes in rice and predicted the potential function of the selected wheat PIP genes.

**Blast using predicted wheat aquaporin protein sequences**

The putative protein products of the full length TaPIP2 and TaPIP1 obtained from the root cDNA in this study was utilised to blastp search against Arabidopsis thaliana and rice database in NCBI (http://blast.ncbi.nlm.nih.gov/) accessed June-July 2009. The resulting matches were aligned using clustalW and a bootstrapped phylogenetic tree was created in MEGA4 (http://www.megasoftware.net/mega4/mega.html), (figure 2.12 and 2.14) using UPGMA method, 1000 replicates.
2.2.5.2 \textit{In silico} search methods to identify expression pattern of rice aquaporins

The full length wheat ESTs CK209661 for \textit{TaPIP2-8} and CK163151 for \textit{TaPIP1-5} were subjected to TIGR Rice database blast search at (http://rice.plantbiology.msu.edu/), accessed June 10, 09). The highest identity match LOC\_Os09g36930, \textit{OsPIP2-7}, putative expressed for CK209661 (EST for \textit{TaPIP2-8}) and LOC\_Os04g47220, \textit{OsPIP1-2} putative expressed for CK163151 (EST for \textit{TaPIP1-5}) were subjected to Rice microarray database (http://www.ricechip.org/) and the respective probe ID for CK209661 (Os.14256.1.S1\_a\_at.) and probe ID for CK163151 (Os.11386.1.S1\_a\_at.). The rice AQP sequence, locus ID and respective probe ID was obtained from (http://www.ricechip.org/cgi-bin/searchrice5.pl?query=Aquaporin, accessed on June, 2009) were submitted in the GeneinvestigatorV3 (expression database) (https://www.genevestigator.com/gv/index.jsp) database. For Example developmental output for the probe Os.11386.1.S1\_a\_at is shown in the appendix 5.

2.2.6 Primer design

NetPrimer (http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html) and Gene Runner Version 3.01 were used to design primers for Polymerase chain reaction (PCR) with the following criteria: between 20 and 40bp long; difference in annealing temperature 5°C to 10 °C between sense and antisense primers; %GC content of ~50%; ends with T or C at the 3’ end (Tendency of Taq polymerase to add A to 3’ end of amplicon); ends with A or G at the 3’ end and primers were checked for minimal secondary structures such as hairpins loop and primer dimers at the 3’ end and no hairpin loop and primer dimers at the 3’ end to make sure both the sense and antisense primers were not self-complementary (Rybicki, 2001).
2.2.6.1 Design of primers to amplify full length PIP genes from wheat

The sequences of primers for PCR amplification of full-length cDNAs and gDNA of TaPIP genes from wheat (T.aestivum L) are listed in (Table 2.5). Primers were designed to amplify two PIP genes, TaPIP2-8 and TaPIP1-5. The partial predicted cds of TaPIP2-8 (Genebank accessions EU177561) and TaPIP1-5 (Genebank accessions EU177550) were used as a query and blastn search was performed against TIGR TAs (Transcript Assemblies) (http://blast.jcvi.org/euk-blast/plantta_blast.cgi) database, to identify the closely related wheat TA sequence. Primer quality was assessed using Net Primer or Gene runner (see above), and the primers (Table 2.5) were designed to amplify full length wheat PIP genes from root and shoot (Gan et al., 2011).

2.2.7 Molecular methods

2.2.7.1 Genomic DNA purification

The frozen wheat leaves (from plants grown conditions, section 2.2.4) were ground into powder with a mortar and pestle in liquid nitrogen. Genomic DNA (gDNA) from leaf tissue was extracted using the Wizard Genomic DNA Purification Kit (Promega). The quality of the purified gDNA was tested by running 5 μL of the total of 100 μL of the purified DNA on a 1.0 % agarose gel in TAE buffer (section 2.2.7.2.2).

2.2.7.2 Quantification of DNA

2.2.7.2.1 Spectrophotometric quantification

The concentration of purified DNA (plasmid or gDNA) was determined using GeneQuant Pro Spectrophotometer (GE Healthcare). Purified DNA was diluted, usually 1:20 or 1:50, with dH2O and absorbance readings were recorded at 230nm, 260nm and 280nm. The concentration of each sample was reported based on 1A260 = 50 μg/mL of double stranded DNA (Sambrook and Russell, 2001). Absorbance
Chapter 2    Cloning and sequencing of wheat PIP genes

ratio of (A260/A280) of DNA was routinely obtained of 1.8 for pure DNA (Sambrook and Russell, 2001)).

2.2.7.2.2  **Agarose gel electrophoresis**

DNA fragment generated by PCR and restriction digests were analyzed using Agarose gel electrophoresis. Agarose gels were prepared at 1% concentrations containing 0.5 μg/mL ethidium bromide (Sambrook and Russell, 2001) in 1X TAE buffer, using Bio-Rad Mini-Sub electrophoresis apparatus, small or large gel casting trays. Electrophoresis was routinely performed at 80-100V for 40-90 minutes. PCR product was analysed by loading 5 μL of DNA mixed with 1 μL of 6X xylene cyanol or bromophenol blue loading dye. Appropriate molecular weight markers were loaded (containing 0.5 μg of DNA) including DNA-GeneRuler DNA Ladder Mix (100 bp–10 kb; Fermentas), or 1kb DNA Ladder (200bp -12 kb; Invitrogen), for estimating the size of the DNA bands (PCR products and restriction digested products). Fluorescent gels were photographed using Digi Doc-It software (Ultra-Violet Products Ltd.), or Chemidoc XRS Documentation Station (Bio-Rad) using Quantity One software (Bio-Rad). The marker GeneRuler DNA Ladder mix was not only used to estimate the size of the DNA bands but also to approximately quantify the DNA in the sample. This was done by comparing the intensity of DNA sample on an agarose gel with that of known concentration of the DNA standards (marker above).

2.2.8  **Polymerase Chain Reaction (PCR)**

2.2.8.1  **PCR conditions**

PCR reactions were generally performed using 100 ng of gDNA template, 12.5μL of 2X Biomix (containing *Taq* polymerase, dNTPs, MgCl₂) and 10 μM of forward and reverse primers were mixed and diluted to 25 μL with sterile dH₂O in 0.2 ml PCR tubes. PCR was carried out using a MyCycler Thermal Cycler (Bio-Rad) with initial denaturation at 95°C for 5 min. Amplification were carried out for 30 cycles of 95°C denaturing for 45 s, annealing temperatures specific for the primers for 45
Chapter 2 Cloning and sequencing of wheat PIP genes

s, and extension at 72°C for 45 s, and the final extension at 72°C for 10 min. PCR products were tested for the presence of the insert by electrophoresing 5 μL of PCR product through agarose gel (section 2.2.7.2.2).

2.2.8.2 Primer synthesis

Primers used for PCR reactions were synthesized by Invitrogen and Sigma and were received as dried pellets. Primers were resuspended in dH₂O to a stock concentration of 100 μM and working concentration of 10μM, diluted from the stock using dH₂O under sterile conditions and stored at -20°C.

2.2.8.3 Amplification of full length PIP genes of common wheat

Different gene sections were amplified from the gDNA and cDNA of common wheat cv. Cranbrook by PCR using gene specific primers (Table 2.5), in 50 μL final volumes. PCRs were performed using typical cycling conditions (section 2.2.8.1), and annealing at 60-64ºC. The results of the PCR products (5 μL) were verified by electrophoresing on 1.0% (w/v) agarose gels to check the results.

2.2.9 Gel purification

2.2.9.1 DNA bands purified from agarose gel

The PCR product (45 μL) was mixed with 9 μL of 6X xylene cyanol (Bromophenol blue was used in few experiments instead of xylene cyanol) loading dye, loaded onto a 1.0% agarose gel and electrophoresed in TAE buffer. The excised DNA bands were purified using the Perfectprep® Gel Cleanup kit and its protocol (Eppendorf). Gel purified product was analysed by loading 1 μL of purified DNA mixed with 4 μL of sterile ddH₂O and 1 μL of 6X xylene cyanol or bromophenol blue loading dye on 1% agarose gel. Purified DNA concentration was estimated by comparing the intensity of the DNA band in the sample after
Chapter 2 Cloning and sequencing of wheat PIP genes

electrophoresis to a standard marker GeneRuler DNA Ladder Mix (100 bp–10 kb; Fermentas) with known DNA concentrations.

2.2.10 Cloning of PCR products

2.2.10.1 Ligation

Gel-purified DNA (section 2.2.9), was ligated with the Promega pGEM-T Easy cloning vector (Promega). Ligations were transformed into chemically competent JM109 E.coli. The Manganese chloride method (protocol obtained from http://microbiology.ucdavis.edu/heyer/protocols/TrafoColi.pdf (Inoue et al. 1990 Gene 96, 23-28)) was used to produce competent JM109 E.coli. Ligation reaction was incubated overnight at 4°C.

2.2.10.2 Transformation

The ligated reaction was transformed into competent JM109 E. coli using a method adapted from the Promega Product Information Sheet #TB095. Competent E. coli were thawed on ice for 5 minutes and 100 μL of competent JM109 E.coli cells were transferred to 1.5 mL microcentrifuge tubes. Up to 4 μL of ligation reaction was combined with the 100 μL of competent cells, and incubated on ice for 20 minutes. The cells were then “heat-shocked” by incubation in a 42°C water bath for 50 seconds, and then re-incubated on ice for 2 minutes. A 0.9 mL aliquot of ice-cold SOC broth (20 g/L Bacto®-tryptone, 5 g/L Bacto®-yeast, 100 mM NaCl2, 25 mM KCl, 200mM Mg2+ stock (203.3 g/L MgCl2.6H2O, 246.5 g/L MgSO4.7H2O), 200 mM glucose) was added to the cells followed by incubation at 37°C for 90 minutes in a shaker ~150 rpm. The reaction was then pelleted by centrifugation at 1000 × g for 10 minutes, resuspended in 200μl of SOC medium, and 100μl plated on each of two LB/Agar plates (10 g/L Bacto®-tryptone, 5 g/L Bacto®-yeast, 5 g/LNaCl2, 15 g/L Bacto®-Agar, pH 7.0) with 100 μg/mL Ampicillin, and incubated overnight at 37°C for 16 hours. For transformation involving blue (non-recombinant) and white (recombinant) screening, the LB/plates
Chapter 2  Cloning and sequencing of wheat PIP genes were supplemented with 0.5 mM IPTG and 80 μg/mL of X-gal (5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside) (Sigma).

2.2.11 Colony PCR

Colony PCR was performed for at least three white colonies from each plate so that a total of six white colonies were screened for the presence of insert in a single ligation reaction. Transformed *E. coli* colonies were individually selected from the LB/Agar plates and suspended in 10 μL of sterile ddH₂O. Selected white colonies were labeled (in numbers) on the plates for future use. Colony PCR was done by adding 1 μL of template (colony dispersed in 10 μL of sterile ddH₂O), 12.5 μL of 2X Biomix (containing *Taq* polymerase, dNTPs, MgCl₂) and 10 μM of each forward and reverse primer, and all were mixed and diluted to 25 μL with sterile ddH₂O. PCR was carried out using a MyCycler Thermal Cycler (Bio-Rad) refer section 2.2.6.1 for cycling conditions. The success of the transformation reaction was determined by electrophoresing 5 μL of colony PCR product through 1% agarose gel (section 2.2.7.2.2).

2.2.12 Plasmid DNA purification

The colonies with the expected product size in the colony PCR were then selected from the LB/Agar plates and inoculated into 5 mL LB broths containing 50 μg/μL of ampicillin and incubated overnight at 37°C with agitation. Plasmid DNA was isolated using the methods described in Wizard® Plus SV Minipreps DNA Purification System (Promega). The success of each purification reactions was assessed by electrophoresing 1 μL of plasmid DNA mixed with 4 μL of sterile ddH₂O and 1 μL of 6X xylene cyanol or Bromophenol blue on a 1.0% (w/v) agarose gels. The concentration of each purified plasmid DNA was determined using UV (Ultraviolet) spectrophotometry (section 2.2.7.2.1) and marker using known DNA concentration (section 2.2.7.2.2). Purified plasmids were stored at -20°C.
Chapter 2 Cloning and sequencing of wheat PIP genes

2.2.13 Restriction digestions

The plasmid DNA was restriction enzyme (RE) digested using the RE EcoRI (Promega) to test whether recombinant plasmids contained the insert of appropriate size. The RE EcoRI, recognition sites flanking the insert in pGEM-T Easy cloning vector (Promega). Method was adapted from Promega usage information sheet, where total volume of 20 μL reaction was set up. Alternatively, all reagent volumes were halved for use in a 10 μL total volume followed by incubation at 37°C for an hour. Digested products were electrophoresed on agarose gels to assess the presence of insert.

2.2.14 DNA sequencing and sequence analysis

2.2.14.1 Purified plasmid DNA for DNA sequencing

Purified plasmid DNA samples were sent to AGD (Applied Genetics Diagnostics, A division of the department of Pathology, The University of Melbourne; [http://agd.path.unimelb.edu.au/electrophoresis-only/index.html](http://agd.path.unimelb.edu.au/electrophoresis-only/index.html)) for sequencing reaction set up and electrophoresis. The purified DNA template and sequencing primer (reverse primer SP6: 5’-ATTTAGGTGACACTATAGAATAC -3’ forward primer T7: 5’-GTAATACGACTCACTATAGGGC -3’) were supplied in separate 1.5ml tubes. The sequencing reaction mixture used at AGD was 1.0 μL of Big Dye Terminator (BDT) version 3.1, 5 μL of 5X BDT Dilution Buffer, 1.0 μL of primer (5 μM), not more than 10 μL DNA (100ng/kb) and the mixture was diluted to 20 μL with dH2O. The cycling conditions performed at AGD was 94°C for 5 minute, then subjected to 30 cycles of 94°C for 10 seconds, 50°C for 5 seconds and 60°C for 3 min ([http://agd.path.unimelb.edu.au/forms/sequencing-reaction-protocol.html](http://agd.path.unimelb.edu.au/forms/sequencing-reaction-protocol.html)). The plasmid DNA was sent to AGRF (Melbourne) (or) AGD (Melbourne) for electrophoresis and capillary separation using a 3730xl DNA Analyzer (Applied Biosystems; California, USA). Each read (sequence obtained) was approximately 800 bp in length on average.
2.2.14.2 DNA sequence analysis

Chromatogram peaks were examined in the BioEdit Sequence Alignment Editor version 7.0.0 (Hall, 1999; http://www.mbio.ncsu.edu/BioEdit/page2.html) to check the quality of sequencing. Sequences were truncated in the places where the separation between peaks was no longer evident. Plasmid DNA’s were generally sequenced in both the directions using the T7(forward) and SP6 (reverse) primers, and the sequences were used to create a consensus sequence. Chromatograms were also checked for mis-calls (incorrectly labeled nucleotides) or mis-counts (wrong number of nucleotides), by comparing the sequences obtained from all the three different clones. Any unclear peaks were labeled as –N‖ in the DNA sequence.

2.2.14.3 DNA sequencing of clones of PIP genes from genomic DNA and cDNA

Purified plasmid DNA containing PIP PCR products (section 2.2.9) were sequenced using sequencing primers in both the forward (T7 primer) and reverse (SP6 primer) directions. At least three clones from each ligation were sequenced in both directions and then compared to create a consensus sequence.

2.2.15 RNA Purification

2.2.15.1 Removal of RNase from materials

Common sources of RNase contamination are hands (skin) and bacteria present on airborne dust particles or laboratory glassware. Therefore fresh gloves were worn at all times to prevent contamination from the above sources. Only RNase-free sterile, disposable plasticware was used for handling RNA. Non-disposable glassware and plasticware were incubated with RNase OUT (G-Biosciences) for 10 seconds and then washed thrice with Milli-Q dH2O, benches and laminar hood was
cleaned with RNase wipes, before use to ensure that it was RNase-free. All buffers and solutions were prepared using DEPC (diethylpyrocarbonate) treated dH2O (Sambrook and Russell, 2001).

### 2.2.15.2 Total RNA extraction from wheat shoot (as well as from leaves earlier) and root tissues.

Wheat plants (Section 2.2.2) were used for RNA extraction. Total RNA was extracted using TRIsure (Bioline) according to the manufacturer’s protocol. In the final step the air-dried pellet was dissolved in 35 μL of DEPC-treated dH2O. Gel electrophoresis was employed to confirm the integrity of total RNA extraction with 2 μL of RNA was loaded on to the 1% agarose gel (section 2.2.7.2.2).

### 2.2.15.3 DNase treatment of total RNA

Total RNA, 33 μL (above) was treated with RQ1 RNase-free DNase I (Promega) and the total volume was scaled up to 50 μL. The mixture was incubated at 37ºC for 30 minutes according to manufacturer instructions (Promega). The RNA was re-purified using LiCl precipitation according to Ambion Technical Bulletin #160 (http://www.ambion.com/techlib/tb/tb_160.html; accessed January 2008). The air-dried pellet was resuspended in 20 μL of DEPC-treated dH2O.

### 2.2.15.4 Agarose gel electrophoresis of RNA samples

Agarose gel electrophoresis was used to assess the quality (integrity and degradation). LiCl-purified total RNA (2 μL), were mixed with 1 μL of 10X RNA loading dye, 3 μL of DEPC-treated dH2O and electrophoresed through a 1.0% (w/v) agarose gel.

### 2.2.15.5 Spectrophotometric Quantification of RNA

The concentration of purified RNA was determined using GeneQuant Pro Spectrophotometer (GE Healthcare). Purified DNA was diluted, usually 1:20 or 1:50, DEPC-treated dH2O and absorbance
readings were recorded at 230nm, 260nm and 280nm. The concentration and purity of each sample was reported based on $1A_{260} = 40 \mu$g/mL of RNA (Sambrook and Russell, 2001). Absorbance ratio of RNA to protein ($A_{260}/A_{280}$) 2.0 was desirable for pure RNA (Sambrook and Russell, 2001).

### 2.2.16 cDNA synthesis

Positive strand of cDNA (first strand) was synthesised using reverse transcriptase (Bioscript; Bioline), followed manufacturer’s protocol, and using oligo dT primers (18T) (Table 2.3) from Invitrogen. A full length open reading frame for the AQP specific gene was not amplified from wheat cDNA prepared using oligo dT primers. Therefore anchored Oligo dT primer (Table 2.3) was used to increase the efficiency of cDNA synthesis from the total RNA extracted using TRIsure (Bioline). Later a tagged anchored oligo dT primer (Table 2.3) was designed to increase the efficiency of cDNA synthesis. Care was taken to design the tag sequence (so that it does not match any sequence in the plant database). The reverse complement of the tagged sequence in the tagged anchored oligo dT primer was used as a reverse primer with TaPIP as a forward primer to capture the 3‘end.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ – 3’)</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligo dT</td>
<td>TTTTTTTTTTTTTTTTTTTT</td>
<td>Reverse</td>
</tr>
<tr>
<td>anchored oligo dT</td>
<td>TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTVN</td>
<td>Reverse</td>
</tr>
<tr>
<td>Tagged anchored</td>
<td>GCTCGCCCTCAGGAACAAATGC</td>
<td>Reverse</td>
</tr>
<tr>
<td>Oligo dT</td>
<td>TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTVN</td>
<td>Reverse</td>
</tr>
</tbody>
</table>

Table 2.3 Primers used to synthesize cDNA

### 2.2.17 cDNA quality assessment

The quality and purity (contamination by gDNA) of cDNA was checked by PCR using 1.0 μL (out of 20 μL) of cDNA in a total volume of 25 μL, using PCR conditions (section 2.2.8.1). The primers used were exon based primers (PIPF1 and PIPR1; see Table 2.4), used to amplify from both cDNA and gDNA of *T. aestivum* cv. Cranbrook. The intron
Chapter 2  

Cloning and sequencing of wheat PIP genes based primers (PIP1AF and PIP1AR; Table 2.4) were used to amplify only from gDNA and not from cDNA, to test the contamination of gDNA in the cDNA preparation. Housekeeping genes actin (Actin1F and Actin1R; Actin2F and Actin2R; Table 2.4) were used to test the quality of 5′-end and 3′-end of cDNA preparation. The cDNA was synthesized for analysis of PIP sequences in leaf and root tissue (21-day-old plant); shoot and root tissue (8-day-old plant).

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ – 3’)</th>
<th>Tm (°C)</th>
<th>PCR product size (bp)</th>
<th>Quality assessment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIPF1</td>
<td>GCCACBYTSCTCTCTCTCTAC</td>
<td>52</td>
<td>Nearly 700</td>
<td>Exon based primers</td>
<td>(Forrest and Bhave, 2008)</td>
</tr>
<tr>
<td>PIPR1</td>
<td>GVCCVACCAGAAGATCC</td>
<td>54.5</td>
<td>889</td>
<td>Intron based primers</td>
<td></td>
</tr>
<tr>
<td>PIPAF</td>
<td>CGTTCCTTGCTCTGCTTG</td>
<td>54.5</td>
<td>889</td>
<td>Intron based primers</td>
<td></td>
</tr>
<tr>
<td>PIPAR</td>
<td>GAGTAGTTATGCGTCCCATCG</td>
<td>54.5</td>
<td>889</td>
<td>Intron based primers</td>
<td></td>
</tr>
<tr>
<td>Actin 1F</td>
<td>GCCACACTGTTCCAATCTATGA</td>
<td>61</td>
<td>369</td>
<td>At the 5’ end</td>
<td>(Mackintosh et al., 2007).</td>
</tr>
<tr>
<td>Actin 1R</td>
<td>TGATGGAATTGTATGTCGCTTC</td>
<td>61</td>
<td>369</td>
<td>At the 5’ end</td>
<td>(Mackintosh et al., 2007).</td>
</tr>
<tr>
<td>Actin 2F</td>
<td>GTTTCCTGGAATTGCTGATCGC</td>
<td>62</td>
<td>410</td>
<td>At the 3’ end</td>
<td>(Yang et al., 2007).</td>
</tr>
<tr>
<td>Actin 2R</td>
<td>ATCATTATTCATACAGCAGGCAA</td>
<td>62</td>
<td>410</td>
<td>At the 3’ end</td>
<td>(Yang et al., 2007).</td>
</tr>
</tbody>
</table>

Table 2.4 Primers used to assess the quality of cDNA.

2.2.18 Cloning of full length PIP cDNA sequences

The cDNA was amplified using gene specific primers TaPIP2F1/Ra1 and TaPIP1F/R (Table 2.5). PCR was performed using 1 μL of cDNA template from T. aestivum cv. Cranbrook leaf and root tissues to amplify TaPIP1 and TaPIP2 genes in 25 μL total volumes. TaPIP2 gene was amplified from the cDNA amplified from the cDNA of 21-day old plant and 8-day old plant, TaPIP1 gene was amplified from the cDNA of 8-day old plant. PCR products were gel-purified (section 2.2.9) and cloned into pGEM-T Easy® plasmids (section 2.2.10). Sequencing was carried out as in section 2.2.14 in both the direction to confirm the sequences were of full length, through alignment in Bioedit
Chapter 2  
Cloning and sequencing of wheat PIP genes 

with wheat partial predicted cDNA, partial gDNA and expressed (EST or TA) sequences identified earlier (data not shown).

2.2.19 Cloning of full length PIP genomic DNA sequences

The gDNA was amplified using gene specific primers TaPIP2 F1/Ra1; TaPIP1F/R (Table 2.5). PCR was performed as in section 2.2.8.1. PCR products were gel-purified (section 2.2.9) and cloned into pGEM®-T Easy plasmids (section 2.2.10). Sequencing was carried out as in section 2.2.14 in both the direction to confirm the sequences were of full length, through alignment in Bioedit with wheat partial predicted cDNA, partial gDNA and expressed (EST or TA) identified earlier. The intron and exon structure of the TaPIP2-8 and TaPIP1-5 was analyzed by aligning the genomic and cDNA sequences obtained for both the genes.

2.2.20 Analysis of wheat PIP gene sequences

The gDNA and cDNA sequences of cloned wheat genes TaPIP2 and TaPIP1 were subjected to a preliminary six-frame translation in Bioedit (http://www.mbio.ncsu.edu/bioedit/bioedit.html) to determine whether or not they were potential MIPs, as well as a blastn (http://blast.ncbi.nlm.nih.gov/) and blastp search (http://blast.ncbi.nlm.nih.gov/) was performed for the gDNA sequence, full length coding sequence and the deduced amino acid sequence. The positions of introns, exons and splice sites in the wheat PIP genes (TaPIP1 and TaPIP2) were analysed by their manual alignment in Bioedit with the wheat PIP gDNA and coding sequences amplified and identified from the root cDNA. The wheat PIP cds (TaPIP1 and TaPIP2) were then translated into putative proteins using Bioedit and aligned with Rice and Arabidopsis amino acid sequences and phylogenetic tree was created.
Chapter 2  Cloning and sequencing of wheat PIP genes

2.2.21 Analysis of the protein structures and features of wheat PIP genes

The putative protein products of (TaPIP1 and TaPIP2), including alternative splice forms of TaPIP2 gene were analysed through ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and cladogram analyses in TreeView to determine whether they belonged to the MIP superfamily, and if so, to which subfamily and group. Information of the wheat putative protein sequences including the predicted molecular weights, pI were determined by inserting the sequence into the ExPASy (http://web.expasy.org/compute_pi/). The number of trans-membrane helices was determined by subjecting each of the wheat protein sequences (TaPIP1 and TaPIP2) into the TMHMM TMpred Server (http://www.cbs.dtu.dk/services/TMHMM/, accessed February 2011). As one of the genes had alternative splice forms, the number of NPA motifs and transmembrane helices were further used for functional MIP analysis. The predicted substrate selectivity (Wallace and Roberts, 2004) were identified manually from ClustalW alignments of putative protein sequences.

2.2.22 Identification of the Orthologous rice and Barley gene for TaPIP1 and TaPIP2

The phylogenetic tree created using putative wheat aquaporin protein sequences with rice, barley, wheat and A.thaliana (Figure 2.13) revealed the candidate rice and Barley aquaporin genes closest to the wheat AQP genes to predict the function of the gene.

2.2.23 Analysis of gene structure of the full length wheat AQP genes TaPIP1 and TaPIP2

The TaPIP1 and TaPIP2 genes obtained from the genomic DNA and the root cDNA were aligned using ClustalW and the intron-exon structures were analyzed. The intron regions were denoted by gaps (-) in the cDNA sequences (Appendix 1 and 2).
Chapter 2  
Cloning and sequencing of wheat PIP genes

2.3 Results and Discussion

2.3.1 Attempts made to amplify full length TaPIP2 and TaPIP1 gene

RNA was extracted from 22 day-old wheat plant. Attempts failed to amplify full length TaPIP2 gene from the 22 day old root and leaf cDNA using primer pairs TaPIP2F1/R1 (see figure 2.5 for primer position) at the temperatures ranging from 60°C to 65°C, whereas this primer pairs amplified the gene from the genomic DNA. Internal reverse middle primer (TaPIP2Rm) was designed based on the ESTs DQ867078 to test the gene for its expression status in the cDNA. TaPIP2Rm yielded a band of size 544bp in the root cDNA (see figure 2.2 A) and not in the shoot cDNA. The result shows that the forward primer and the internal reverse primer are in the expressed sequence but a reverse primer had to be found to capture the full length coding sequence of TaPIP2 gene in order to study the function of the protein. The EST (DQ867078) used to design the primers TaPIP2F1/R1 was subjected to ORF finder to predict the potential ORF to code for the functional protein. Based on the obtained results a new reverse primer (R2) was designed to amplify full length TaPIP2-8 from both root and shoot cDNA. The primer pair TaPIP2F1/R2 did not yield any bands in the root and shoot cDNA and was not attempted at genomic level. Therefore the final primer pair TaPIP2F1/ Ra1Ra2 was designed using the EST CK209661, obtained from TA search. Nested PCR reaction was performed and yielded products in both gDNA (Figure 2.2 B) and in the cDNA (Figure 2.4 A & B) of root and not in the shoot. The primer pairs TaPIP2F1/R1 and TaPIP2 F1/Ra1 yielded the same sequence at the genomic level. The difference between the reverse primers R1, Ra1 and Ra2 was their positions, R1 is located before the stop codon, Ra1 on the 3’UTR region and Ra2 before the 3’UTR/on the stop codon (Figure 2.5).
## 2.3.2 cDNA quality testing

Total RNA was purified from 8-day old shoot and root of the plant *T. aestivum* cv. Cranbrook and cDNA was reverse transcribed, and the quality of the cDNA preparation tested for the quantity of cDNA and lack of genomic DNA contamination. The cDNA was confirmed suitable for further analysis if intron-based primers (PIPAF and PIPAR, used from Forrest & Bhave) resulted in lack of amplification, while the exon-based primers (PIPF1 and PIPR1, used from Forrest and Bhave, 2008) resulted in strong amplification of a single band of the appropriate size of 600bp (see Figure 2.1, B). The quality of the 5' and 3' end was also tested using the primers Actin1F and Actin1R (5' end), 369 bp and Actin2F and Actin2R (3' end), 410 bp (see Figure 2.1 A). The cDNA was then used to amplify wheat PIP genes using gene specific primers (Table 2.5) and the PCR products were cloned and sequenced.

### Table 2.5 Primers designed to amplify full length wheat PIP genes for this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ – 3’)</th>
<th>Direction</th>
<th>Target gene</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIP2 F1</td>
<td>AAAggatccATGGCCAAGGAGGTAGCGGAGG</td>
<td>Forward</td>
<td>PIP2</td>
<td>Full-length amplification from genomic DNA</td>
</tr>
<tr>
<td>PIP2 R1</td>
<td>AAAggatccTCAGTTGCTCCGGCTTGCTCC</td>
<td>Reverse</td>
<td>PIP2</td>
<td></td>
</tr>
<tr>
<td>PIP2 Rm</td>
<td>CGGTGTAAGGACACGGAAGG</td>
<td>Reverse</td>
<td>PIP2</td>
<td></td>
</tr>
<tr>
<td>PIP2 cRm</td>
<td>GCCACATCTGCTCGT</td>
<td>Reverse</td>
<td>PIP2</td>
<td></td>
</tr>
<tr>
<td>PIP2 R2</td>
<td>AAgaattcTCATGCATATAAGGTTAG</td>
<td>Reverse</td>
<td>PIP2</td>
<td></td>
</tr>
<tr>
<td>PIP1 F</td>
<td>AAAggatccATGGGAAAGGAGGAGGAC</td>
<td>Forward</td>
<td>PIP1</td>
<td></td>
</tr>
<tr>
<td>PIP1 R</td>
<td>AAgaattcTTAGGACTGTTCTGAACGGGGATCGC</td>
<td>Reverse</td>
<td>PIP1</td>
<td></td>
</tr>
<tr>
<td>PIP2 Fab</td>
<td>AAgaattcATGGCCAAGGAGTGGAGCGA</td>
<td>Forward</td>
<td>PIP2</td>
<td></td>
</tr>
<tr>
<td>PIP2 Ra2</td>
<td>AAgaattcTTAATTATATGGCCGGCCCCG</td>
<td>Reverse</td>
<td>PIP2</td>
<td></td>
</tr>
<tr>
<td>PIP2 Ra1 (Nested PCR)</td>
<td>AAgaattcAGGCCGTTGCGGTACCGG</td>
<td>Reverse</td>
<td>PIP2</td>
<td></td>
</tr>
</tbody>
</table>

Restriction endonuclease underlined BamH1 and restriction endonuclease highlighted in ash colour EcoR1.
Chapter 2  Cloning and sequencing of wheat PIP genes

Figure 2.1 Example of quality test of cDNA preparations

**A:** PCR test used to determine the success and quality of 5’ and 3’ end of shoot cDNA synthesised from Early Seedling Stage plants. The 396 bp band in the molecular weight marker is indicated. Lanes **M:** 1kb Ladder; 1-4: shoot cDNA, 369 bp (Actin1- Housekeeping gene primers, to check the quality of 5’ end of the cDNA); 5: negative control (No Reverse transcriptase); 6-9: shoot cDNA, 410 bp (Actin2- 3’ Housekeeping gene primers, to check the quality of 3’ end of the cDNA); 10: negative control (No Reverse transcriptase).

**B:** PCR test used to determine the quality of cDNA synthesis, whether the cDNA is free from contaminating gDNA. The 506 bp band in the molecular weight marker is indicated. Lanes **M:** 1kb DNA Ladder Mix; 1-4: shoot cDNA, 600 bp (PIPF1/R1 exon primers); 5: negative control (No Reverse transcriptase); 6-9: shoot cDNA, 410 bp (PIP1AF/R intron primers); 10: negative control (No Reverse transcriptase).

### 2.3.3 Isolation of two full length aquaporin genes from wheat

This study mainly concentrated on amplifying full length coding sequence to express in baculovirus and test the functional water transport activity using a newly developed method. However, the respective genomic DNA sequences were amplified to verify the newly designed gene specific primers, to optimize the annealing temperature and to derive the gene structure of the amplified coding sequences. The two full length aquaporin genes attempted to amplify in this study are *TaPIP2-8* and *TaPIP1-5* (according to the classification of Forrest and Bhave). These genes are referred as *TaPIP2* and *TaPIP1* in this chapter.
Amplification of those two genes from genomic DNA resulted in two variants of TaPIP1 and TaPIP2 after screening 10 different colonies. The primers used to amplify the full length genes were designed using the TA which showed greatest identity with the TaPIP1-5 (EU177550) and TaPIP2-8 (EU177561) genes. Table 2.6 shows the top six TA matches, where all the matches were verified for the full length coding sequence.

The top match for TaPIP2-8 is comprised of 10 ESTs and all the ESTs were partial in length. The second closest match was CK209661, which was full in length and so the primers TaPIP2F1/ Ra1 Ra2 were designed based on the EST CK209661 to amplify full length TaPIP2-8.

The closest TA match of TaPIP1-5 (EU177550) was TA53971_4565 comprised of 37 ESTs, in which 36 were partial in length and one of the ESTs among 37, CK163151 was full length and therefore the primer pairs TaPIP1F and TaPIP1R were designed based on the EST CK163151 to amplify full length TaPIP1-5.

<table>
<thead>
<tr>
<th>Partial TaPIP2-8 (EU177561)</th>
<th>Partial TaPIP1-5 (EU177550)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA52348_4565 (had 10 ESTs)</td>
<td>TA53971_4565 (had 37 ESTs)</td>
</tr>
<tr>
<td>CK209661</td>
<td>CK210483</td>
</tr>
<tr>
<td>CK207081</td>
<td>TA53998_4565</td>
</tr>
<tr>
<td>CK210019</td>
<td>CK211013</td>
</tr>
<tr>
<td>CK209321</td>
<td>CD871295</td>
</tr>
<tr>
<td>TA51100_4565</td>
<td>TA54443_4565</td>
</tr>
</tbody>
</table>

Table 2.6 TA (Transcript assembly) matches obtained for the partial TaPIP1-5 (EU177550) and TaPIP2-8 (EU177561) genes from http://plantta.jcvi.org/search.shtml

The ESTs CK163151 (expressed in crown and leaf tissues in wheat) and CK209661 (expressed in crown and developmental stages of spike formation in wheat) were used to design primers and amplify TaPIP1-5 and TaPIP2-8 gene from root cDNA.
Chapter 2  Cloning and sequencing of wheat PIP genes

2.3.3.1  Isolation of full length TaPIP2 gene from wheat

2.3.3.1.1  Amplification of TaPIP2 gene from genomic DNA

Full length TaPIP2 gene was amplified from genomic DNA of wheat *T. aestivum* cv. Cranbrook using different primer pairs TaPIP2 F1/R1 (full length gene of gDNA) and TaPIP2 F1/Ra1 (full length gene of gDNA and cDNA) (see figure 2.5 for primer position). These primers amplified products of 962 bp and 1007bp from the genomic DNA. As explained in the section 2.3.1 primer pair TaPIP2F1/R1 amplified a product of 962 bp from gDNA and not from cDNA for some unknown reasons. Therefore the primer pair TaPIP2F1/Ra1 was used to amplify TaPIP2 gene from both cDNA and gDNA. The difference in the primer pairs is the reverse primer and its position (Figure 2.1), R1 primer is located on the stop codon; Ra2 and Ra1 primers are located on the 3‘ UTR (Untranslated region) region. Three clones were screened from the primer pairs (TaPIP2F1/R1 and TaPIP2F1/Ra1) and sequenced in both the directions. The sequence analysis revealed that the two clones from the primer pairs TaPIP2F1/R1 and all the three clones from the primer pairs TaPIP2F1/Ra1 amplified an identical gene and denoted as *TaPIPA* in this study and later subgrouped (see section 2.3.3.1.2). One out of three clones sequenced from the primer pair TaPIP2F1/R1 amplified a homeologous gene (sharing high identity with the sequence and variation in the intron sequence) with some SNPs (single nucleotide polymorphism) in the intron region between 735 and 865 bp, (Figure 2.3). Due to this difference; this homeologous gene obtained from one of the three clones is denoted as *TaPIPB*. Therefore two homeologous genes of TaPIP2 were amplified by screening six clones. The *TaPIPA* and *B* share an identity of 94% at the nucleotide level and 98% in the protein level.
Chapter 2  Cloning and sequencing of wheat PIP genes

Figure: 2.2  PCR with gDNA of TaPIP2 gene with different primer pairs.
A: Lane M - 1Kb DNA Ladder Mix, Lane 1 and 2 - Primer pair PIP2 F1/R1
Lane 3 - Primer pair PIP2 F1/Rm.
B: PCR with gDNA of PIP2 gene with PIP2 F1/Ra1. Lane 1: Primer pair PIP2 F1/Ra1,
Bright band was obtained at 65°C of ~1kb.

2.3.3.1.2  Subgrouping of TaPIP2 gene from genomic DNA

The nucleotide sequence of TaPIPA and B were further compared and aligned with the known partial wheat gDNA, wheat PIPs and TIPs in the database (Accession numbers EU177547 to EU177570) identified in the previous study by (Forrest and Bhave, 2008). The TaPIPA and B genes share highest identity of 95% and 93% with TaPIP2-9 at gDNA level and 98% and 99% identity with TaPIP2-9 at predicted amino acid sequence level (Table 2.7). The obtained full length TaPIPA and B were therefore designated as TaPIP2-9a and TaPIP2-9b because they showed high similarity in the sequences of all exons of previously reported TaPIP2-9 (EU177562) and single intron (Intron III) varied in the length 108, 110 and 107 in TaPIP2-9, TaPIP2-9a and TaPIP2-9b respectively (Table 2.8).

The present work is the first report of the full length genomic sequences of the wheat TaPIP2-9a, a homeolog of previously identified
Chapter 2 Cloning and sequencing of wheat PIP genes

*TaPIP2-9* partial genomic sequence. Similarly a full length genomic sequence of *TaPIP2-9b* gene was also identified in this study, though the partial genomic sequence of *TaPIP2-9b* (intron III of 107bp) was reported by Yousif and Bhave, 2011, which may also represent homeologs of the previously identified genes (Yousif and Bhave, 2011).

<table>
<thead>
<tr>
<th>Genebank accession</th>
<th>Partial sequences</th>
<th>% identity <em>TaPIPA</em></th>
<th>% identity <em>TaPIPB</em></th>
<th>% identity <em>TaPIPA</em></th>
<th>% identity <em>TaPIPB</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>EU177562</td>
<td>TaPIP2-8</td>
<td>89</td>
<td>87</td>
<td>84</td>
<td>84</td>
</tr>
<tr>
<td>EU177562</td>
<td>TaPIP2-9*</td>
<td>95*</td>
<td>93*</td>
<td>99*</td>
<td>98*</td>
</tr>
</tbody>
</table>

Table 2.7 Summary of the % identities of full length *TaPIP2-9* gene obtained from genomic DNA with the partial *TaPIP* genes in the database. The percentage identity value indicated by star represents the highest identity score corresponding to the respective *TaPIP* sequence from the database.
Figure 2.3 Wheat TaPIP2 gDNA sequences. Sequence alignment of cloned wheat PIP2-9a and PIP2-9b from gDNA, prepared using ClustalW in Bioedit. A dot (.) indicates nucleotide conservation at that position with TaPIP2-9a. A dash (-) indicates lack of sequence at that position. Grey shaded region denotes the changes in the nucleotide towards the 3’ end of the gene. The introns are boxed.
2.3.3.1.3 Amplification of TaPIP2 gene from cDNA

Figure 2.4 Amplification of full length TaPIP2 gene from 8-day-old root cDNA

A: lane 1 to 4 shows the Gradient PCR to amplify the full length PIP2 gene using PIP2 F1/Ra1 primers, temperatures ranging from 66°C to 69°C.

B: Nested PCR reaction (second round PCR) to amplify TaPIP2 gene using PIP2 Fab/R2 primers at 66°C.

The full length TaPIP2 cds was amplified from the early seedling stage (8-day-old wheat plant) of root cDNA (see figure 2.4) using the primer pair TaPIP2 F1/Ra1 & Ra2, with a nested PCR reaction, amplification was observed in the shoot cDNA (data not shown). Attempts failed to amplify full length TaPIP2-8 gene from the 21-day old root and leaf cDNA using primer pairs TaPIP2F1/R1 (see section 2.2.6 for primer design) at the temperatures ranging from 60°C to 65°C. Internal reverse middle primer (TaPIP2 Rm) was designed to amplify the partial gene to verify the presence of the expression of the gene. The partial gene was expressed in the root cDNA, and the 544bp fragment (data not shown) was cloned and sequenced. The obtained partial coding sequence was utilized to do further search against TAs and ESTs of Triticum aestivum. The EST CK209661 was 99% identical to the nucleotide sequence of TaPIP2 partial cds (EU177561) and the reverse primers, TaPIP2 Ra1 and Ra2 was designed based on the 3’UTR (see figure 2.5 for primer position) to capture the full length cds and nested PCR was performed. The expression of rice aquaporins were analyzed in the rice micro array database (http://www.ricechip.org/) for the stages of expression. The analysis revealed
Chapter 2  Cloning and sequencing of wheat PIP genes

that most of the aquaporins are expressed at the early seedling stage (8-day old plant). Therefore the total RNA was isolated and cDNA was synthesized from the early seedling stage of shoot and root of wheat plant *T. aestivum* cv. Cranbrook. The PCR product obtained from the primer pair *TaPIP2* F1/Ra1 Ra2 is 897bp (Figure 2.4 A and B). The obtained PCR product was gel purified cloned and three colonies were screened for the presence of the insert. One of the colonies showed a smaller band at ~750bp (data not shown). Subsequently a total of ten colonies were screened to verify the insert. A band at ~750bp appeared for five colonies and the other five colonies showed up a band at ~900bp. Three clones from the two length variants were sequenced and analyzed. The primer positions were located on the amplified full length sequence.

![Figure 2.5 Primer positions of *TaPIP2-9* gene.](image)

Primer positions located on the full length *TaPIP2-9* gene amplified from the genomic and root cDNA of wheat.

2.3.3.1.4  **Analysis of full length cds and alternate splice variant of *TaPIP2* gene**

The alignment of the nucleotide sequence of the larger band (*TaPIP2* cds) and smaller band (*TaPIP2AS*) (see figure 2.6) from the *TaPIP2* gene reveals that the smaller band is an alternate splice variant of the *TaPIP2* gene, missing 156bp in the exon region from 261bp to 416bp (Figure 2.6), which codes for 52 amino acids and encompasses transmembrane helix 2 (TMH2), Loop B bearing first NPA motif and transmembrane helix 3 (TMH3). The larger band 897bp is a full length cds, putative protein sequence, 284 amino acids codes for a functional aquaporin protein, whereas the alternate splice variant codes for 232 amino acids lacking the TMH 2, loop B and TMH 3. Therefore the alternate splice variant cannot form the hourglass structure without TMH 2, loop B and TMH 3, hence is non-functional as an aquaporin.
Cloning and sequencing of wheat PIP genes

**Figure 2.6 Wheat PIP2 cds sequences**

Sequence alignment of cloned PIP2 Coding sequences of wheat, prepared using ClustalW.

*TaPIP2* cds denotes full length coding sequence and *TaPIP2* AS, denotes alternate spliced (AS) sequence obtained from root cDNA. A star (*) indicates nucleotide conservation at that position with *TaPIP2* cds. A dash (-) indicates lack of sequence at that position. The alternate spliced region (lack of exon) is highlighted in black denoted in (-). The arrow at position 735 denotes the intron position. The stop codon is underlined.
2.3.3.1.5 Comparison of TaPIP2 gene from root cDNA with the predicted cds of TaPIP2-9a and TaPIP2-9b obtained from genomic DNA

The full length cds and alternate splice variant of TaPIP2 gene obtained from root cDNA were compared with the predicted partial coding sequence (cds) of TaPIP2-9a and TaPIP2-9b obtained from the genomic DNA. The full length TaPIP2 cds (897 bp) shares 99% and 96% identity with TaPIP2-9a and TaPIP2-9b. The TaPIP2 cds is designated as TaPIP2-9a cds because of the highly conserved exons and also the predicted amino acid sequence shares highest identity of 99% with the TaPIP2-9a gDNA. The alternate splice (AS) variant of TaPIP2 cds (727 bp) shares 95% and 93% identity with predicted cds of TaPIP2-9a and TaPIP2-9b. Therefore the alternate splice (AS) variant of TaPIP2 cds (727 bp) is designated as TaPIP2-9aAS.

2.3.3.1.6 Gene structure of TaPIP2-9 gene

By comparing the exon–intron regions from genomic and cDNA clones of wheat aquaporin genes (Appendix 2), several structural features of interest were observed. TaPIP2-9a and TaPIP2-9b showed sequence homology of (94%) with identical exon-intron structures and lengths (Table 2.10). The gene structure exhibited one intron and two exons. Two NPA motifs were found in the first exon (Figure 2.7). The TaPIP2-9a and TaPIP2-9b shares 73% identity in the intron region with 30 SNPs and 96% and 99% identity in the exon 1 and exon 2 region. Exons were well conserved within each aquaporin subgroup compared to introns. The two isoforms of this PIP could be present in the two different genomes of the hexaploid wheat.
Figure 2.7 Gene structure of TaPIP2-9 gene.

**A (1) and A (2):** Gene structure of TaPIP2-9a (965 bp) and TaPIP2-9b (962 bp). The boxed region matches the known partial TaPIP2-9a cds (accession number EU177561) and the unboxed region was amplified using gene specific primers to capture the N and C terminal region. **B and C:** Gene structure of TaPIP2-9a cds and alternatively spliced TaPIP2-9 gene obtained from root cDNA. The dark black line in the exon 1 denotes the two NPA motifs and the ash box denotes the region of the missing 156 bp.
### Table 2.8 % identity of introns and exons of TaPIP2-9 gene isolated from gDNA and cDNA of wheat.

The percentage value denotes the %identity between the introns of the homeologous genes TaPIP2-9a and TaPIP2-9b and also the %identity between the exons of TaPIP2-9a cds and the predicted exons from the genomic DNA (right). Thick black line separates the introns (left) and exons (right) of TaPIP2-9. “—“no intron 1 and intron 2.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Intron III (bp)</th>
<th>Exon I (bp)</th>
<th>Exon II (bp)</th>
<th>% identity of predicted coding sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP2-9a</td>
<td>110</td>
<td>735</td>
<td>120</td>
<td>73%</td>
</tr>
<tr>
<td>TaPIP2-9b</td>
<td>107</td>
<td>735</td>
<td>120</td>
<td>96%</td>
</tr>
<tr>
<td>TaPIP2-9a</td>
<td>110</td>
<td>735</td>
<td>120</td>
<td>96%</td>
</tr>
<tr>
<td>TaPIP2-9b</td>
<td>107</td>
<td>735</td>
<td>120</td>
<td>96%</td>
</tr>
<tr>
<td>TaPIP2-9</td>
<td>108</td>
<td>594</td>
<td>19</td>
<td>98%</td>
</tr>
<tr>
<td>TaPIP2-9</td>
<td>108</td>
<td>594</td>
<td>19</td>
<td>98%</td>
</tr>
<tr>
<td>TaPIP2-9a cds</td>
<td>110</td>
<td>-</td>
<td>735</td>
<td>99%</td>
</tr>
<tr>
<td>TaPIP2-9a</td>
<td>-</td>
<td>-</td>
<td>594</td>
<td>99%</td>
</tr>
<tr>
<td>TaPIP2-9a AS</td>
<td>-</td>
<td>-</td>
<td>578</td>
<td>100%</td>
</tr>
<tr>
<td>TaPIP2-9a</td>
<td>107</td>
<td>735</td>
<td>120</td>
<td>95%</td>
</tr>
</tbody>
</table>
2.3.3.2 Isolation of full length *TaPIP1* gene from wheat

2.3.3.2.1 Amplification of *TaPIP1* gene from genomic DNA

Full length *TaPIP1* gene was amplified from genomic DNA of wheat *T. aestivum* cv. Cranbrook using primer pair *TaPIP1F* and *TaPIP1R* (see figure 2.5 for primer position). These primers amplified a product of ~1600 bp from the genomic DNA at 63°C (Figure 2.8A, lane 2). The obtained product was gel purified and cloned. The verification of the gel purified product (Figure 2.8B) in the 1% agarose gel showed that two genes were amplified using the primer pair *TaPIP1F* and *TaPIP1R* at two different sizes 1541 bp and 1643 bp (precise values from later sequencing) from the genomic DNA. The two genes were very close to each other in the gel (Figure 2.9B) and therefore these two bands were gel purified in the same reaction tube and cloned together. To obtain the two genes ten different colonies were screened and restriction digested (Figure 2.8C). Figure 2.8C shows that the five colonies (Lanes 1-4 & 6) have the insert of 1541 bp and the other five colonies (Lanes 5 & 7-10) have the insert of 1642 bp. The two *PIP1* genes amplified from *TaPIP1F* / R primers at genomic level were denoted as *TaPIPC* and *TaPIPd*. The *TaPIPC* and *TaPIPd* genes shares 92% identity at the nucleotide level and 99% at the protein level.
Figure 2.8 PCR to amplify full length TaPIP1 gene from gDNA, gel purified and restriction digested product.
A: Gradient PCR of TaPIP1 gene from gDNA, Lane 1-2 - 60°C – 63°C, amplification observed at 63°C. B: Gel purified product of TaPIP1 gene obtained from genomic DNA of wheat (Cv: Cranbrook). C: EcoRI Restriction digested product of TaPIP1 gene isolates, Lane M- 1Kb ladder, Lanes 1-4 & 6 – 1541 bp; Lanes 5 & 7-10 – 1642 bp.
2.3.3.2.2 Subgrouping of TaPIP1 gene from genomic DNA

The nucleotide sequence of TaPIPC and D were further compared and aligned with the known partial wheat gDNA, wheat PIPs and TIPs in the database (Accession numbers EU177547 to EU177570) identified in the previous study by (Forrest and Bhave, 2008). The TaPIPC and D genes share highest identity of 99% and 98% with TaPIP1-6 and TaPIP1-2 at gDNA level (Table 2.7). The predicted amino acid sequence of TaPIPC gene shares 98% identity with both TaPIP1-5 and TaPIP1-6, TaPIPD gene also shares 98% identity with TaPIP1-2 and TaPIP1-5. The obtained full length TaPIPC and D was designated as TaPIP1-6a and TaPIP1-2a because they showed high similarity in the sequences of all exons to the previously reported partial sequences TaPIP1-6 and TaPIP1-2 (Table 2.10). The intron I and III (85 and 491 bp) in TaPIP1-6a and TaPIP1-6, but varied in the length of intron II (99 and 97 bp) in TaPIP1-6a and TaPIP1-6 respectively (Table 2.10), whereas the introns in the TaPIP1-2a and TaPIP1-2 are of same length in the intron II and III (122 and 552 bp) but varied in the length of intron I (102 and 105 bp). The present work is the first report of the full length genomic sequences of the wheat TaPIP1-2a and TaPIP1-6a, a homeolog of previously identified TaPIP1-2 and TaPIP1-6 partial genomic sequence.

<table>
<thead>
<tr>
<th>Genebank accession</th>
<th>Partial sequences</th>
<th>% identity TaPIPC</th>
<th>% identity TaPIPD</th>
<th>% identity TaPIPC</th>
<th>% identity TaPIPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIP 1</td>
<td>TaPIP1-2</td>
<td>-</td>
<td>98*</td>
<td>-</td>
<td>98*</td>
</tr>
<tr>
<td>EU177550</td>
<td>TaPIP1-5</td>
<td>96</td>
<td>-</td>
<td>98*</td>
<td>98*</td>
</tr>
<tr>
<td>EU177551</td>
<td>TaPIP1-6</td>
<td>99*</td>
<td>-</td>
<td>98*</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.9 Summary of the % identities of full length TaPIP1-5 gene obtained from genomic DNA with the partial TaPIP genes from the database.

The percentage identity value indicated by star (*) represents the highest identity score corresponding to the respective TaPIP sequence from the database.
Chapter 2
Cloning and sequencing of wheat PIP genes

(a) Alignment of genomic sequences of TaPIP1-6a with the closely related TaPIP1-6 (EU177551).

<table>
<thead>
<tr>
<th>Exon 1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-6a gDNA</td>
<td>ATGGAGGGCAAGGAGAGGAGCCTGGGCGCTGGGGCAACAAATGACTTCGAGCGCAGCCCAGTGCCCAAGGGGTGGAGGACAGCAAGACACTACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaPIP1-6 gDNA</td>
<td>110</td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exon 2</th>
<th>310</th>
<th>320</th>
<th>330</th>
<th>340</th>
<th>350</th>
<th>360</th>
<th>370</th>
<th>380</th>
<th>390</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-6a gDNA</td>
<td>TCGCCCTGTACTCTACCCCGCCATCTCCGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaPIP1-6 gDNA</td>
<td>410</td>
<td>420</td>
<td>430</td>
<td>440</td>
<td>450</td>
<td>460</td>
<td>470</td>
<td>480</td>
<td>490</td>
<td>500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Int</th>
<th>510</th>
<th>520</th>
<th>530</th>
<th>540</th>
<th>550</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-6a gDNA</td>
<td>CAGGTACATCGCCACCCCGGGCCGCGCTGGACATCGGAGGGAACCTTCTGGGCGGGGGCATTTGCAATCATCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaPIP1-6 gDNA</td>
<td>560</td>
<td>570</td>
<td>580</td>
<td>590</td>
<td>600</td>
</tr>
</tbody>
</table>
Chapter 2  
Cloning and sequencing of wheat PIP genes
Chapter 2

Cloning and sequencing of wheat PIP genes

TaPIP1-6a gDNA
TaPIP1-6 gDNA

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400

TaPIP1-6a gDNA
TaPIP1-6 gDNA

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500

TaPIP1-6a gDNA
TaPIP1-6 gDNA

1510 1520 1530 1540

TaPIP1-6a gDNA
TaPIP1-6 gDNA

1650 1660 1670 1680 1690 1700 1710 1720 1730 1740

Exon 1

Intron

Hordei

Triticum

Oryza
Chapter 2

Cloning and sequencing of wheat PIP genes

(b) Alignment of genomic sequences of TaPIP1-2a with the closely related TaPIP1-2 (EU177547).
Chapter 2

Cloning and sequencing of wheat PIP genes

TaPIP1-2a gDNA

810 820 830 840 850 860 870 880 890 900

TaPIP1-2 gDNA

910 920 930 940 950 960 970 980 990 1000

Exon 1

Hordes

Intron

Hordes

Intron 2

Hordeum

Intron

Triticum

Intron

Triticum

Intron

Oryza

Exon 1

Intron

Hordeum
**Figure 2.9 Alignment of TaPIP1-2a with TaPIP1-2 (EU177547) and TaPIP1-6a and TaPIP1-6 (EU177551) genomic sequence.**

The regions highlighted in black are intron regions, unshaded regions are exons and the region highlighted in ash denotes the partial gene in the database lacking N and C terminal.

Alignment created in ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and displayed in Bioedit
2.3.3.2.3 Amplification of TaPIP1 gene from cDNA

The full length TaPIP1 cds was amplified from the early seedling stage (8-day-old wheat plant) of root and shoot cDNA (see figure 2.10A and B) using the primer pair TaPIP1F and TaPIP1R (see section 2.2.6 for primer design) from the EST CK163151. The PCR was conducted at the temperatures ranging from 64°C – 67°C for both the root and shoot cDNA, the PCR product obtained at 64°C in the root cDNA (Figure 2.10B, lane 1) was reamplified using the same primer pair TaPIP1F and TaPIP1R at 64°C (Figure 2.10C), for a better yield for successful cloning. The reamplified PCR product increased the amount of DNA and was cloned and sequenced. Three clones were sequenced and analyzed. The PCR product obtained from shoot cDNA (visible at 67°C, Figure 2.10A, lane 4) was very faint and was not cloned and sequenced.

Figure 2.10 PCR to amplify full length TaPIP1 gene from shoot and root cDNA.
A: Gradient PCR of Shoot cDNA to amplify the full length TaPIP1 gene, Lane 1-4 - 64°C – 67°C, amplification observed at 67°C. B: Gradient PCR of Root cDNA to amplify the full length TaPIP1 gene, Lane 1-4 - 64°C – 67°C, amplification observed at all the temperatures but greatest at 64°C. C: Reamplification of TaPIP1 from root cDNA at 62°C with the same primer pairs for 30 cycles, 867 bp.
2.3.3.2.4 Comparison of TaPIP 1 gene from root cDNA with the TaPIP1-6a and TaPIP1-2a obtained from the genomic DNA

The full length cds of TaPIP1 gene were compared with the predicted coding sequence (cds) of TaPIP2-9a and TaPIP2-9b obtained from the genomic DNA. The full length TaPIP1 cds (867 bp) shares 98% identity with TaPIP1-6a and TaPIP1-2a at nucleotide level and 97% and 98% at protein level. The TaPIP1 cds is designated as TaPIP1-2a cds because the predicted amino acid has 98% identity with the TaPIP1-2a.

2.3.3.2.5 Gene structure of TaPIP1-2a and TaPIP1-6a gene

The exon-intron structure of TaPIP1 gene is quite different from the TaPIP2 gene (Refer appendix 1 for alignment). Both the TaPIP1-2a and TaPIP1-6a genes have 3 introns and 4 exons (Figure 2.11A and B). The lengths of the introns are variable in size between the TaPIP1-2a and TaPIP1-6a genes, whereas the exon sizes are identical and highly conserved except for the exon 3 144 and 142bp. The intron 1 of TaPIP1-2a and TaPIP1-6a genes are 85 bp and 102 bp in length and shares 87% identity, intron 2 shares 76% identity and are 99 bp and 122 bp in length, intron 3 shares 85% similarity and are 491 bp and 552 bp in length. The exon 1 and exon 3 of TaPIP1-2a and TaPIP1-6a genes are 335 bp and 141/142 bp and shares 99% identity, exon 2 and exon 4 of TaPIP1-2a and TaPIP1-6a genes are 295 bp and 95 bp sharing 98% identity.
Chapter 2  
Cloning and sequencing of wheat PIP genes

Figure 2.11 Gene structure of TaPIP1 gene
A: Gene structure of TaPIP1-2a, 1541 bp. B: Gene structure of TaPIP1-6a, 1643 bp. C: The gene structure of TaPIP1-5 cds. The boxed regions matches the known partial TaPIP1-2a ds (accession number EU177550) and the unboxed region was amplified using gene specific primers designed from an EST CK163151 to capture the N and C terminal region, the dark black lines in the exon 2 and exon 3 denotes the NPA motifs.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Intron I (bp)</th>
<th>Intron II (bp)</th>
<th>Intron III (bp)</th>
<th>Exon I (bp)</th>
<th>Exon II (bp)</th>
<th>Exon III (bp)</th>
<th>Exon IV (bp)</th>
<th>Predicted amino acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP C</td>
<td>85</td>
<td>99</td>
<td>76</td>
<td>491</td>
<td>85</td>
<td>335</td>
<td>98</td>
<td>141</td>
</tr>
<tr>
<td>TaPIP D</td>
<td>102</td>
<td>122</td>
<td>%</td>
<td>552</td>
<td>%</td>
<td>335</td>
<td>%</td>
<td>142</td>
</tr>
<tr>
<td>TaPIP C</td>
<td>85</td>
<td>99</td>
<td>94</td>
<td>491</td>
<td>99</td>
<td>335</td>
<td>98</td>
<td>295</td>
</tr>
<tr>
<td>TaPIP1-6</td>
<td>85</td>
<td>97</td>
<td>%</td>
<td>491</td>
<td>%</td>
<td>335</td>
<td>98</td>
<td>295</td>
</tr>
<tr>
<td>TaPIP D</td>
<td>102</td>
<td>120</td>
<td>%</td>
<td>552</td>
<td>%</td>
<td>335</td>
<td>98</td>
<td>142</td>
</tr>
<tr>
<td>TaPIP1-2</td>
<td>105</td>
<td>99</td>
<td>100</td>
<td>552</td>
<td>%</td>
<td>335</td>
<td>98</td>
<td>142</td>
</tr>
</tbody>
</table>

Table 2.10 % identity of introns and exons of TaPIP1-5 gene isolated from gDNA of wheat.

The percentage value denotes the %identity between the introns of both the versions of TaPIP1 (left) and also the %identity between the exons of TaPIP1-5 (right). Thick black line separates the introns (left) and exons (right) of TaPIP1. The exon I and exon IV highlighted in black are partial in length in the database.
2.3.4 Comparison of the TaPIP1-2a, TaPIP2-9a and TaPIP2-9a AS coding sequence with the known full length ESTs

The coding sequences of TaPIP1-2a and TaPIP2-9a and the resulting putative protein sequences (nucleotide and protein alignment file in appendix 6 (a) and (b)) were aligned in ClustalW with those of ten complete coding sequences of wheat PIPs mention in chapter 1 (Section 1.8).

TaPIP2-9a cds, TaPIP2-9a AS, TaPIP2-9a and TaPIP2-9b exhibited highest sequence identity of 95%-97% at cDNA level and 97-99% at protein level to DQ867078, a TaPIP2 (AQP4) (Table 2.11). The ClustalW alignment (Table 2.11) showed that TaPIP1-2a cds, TaPIP1-2a and TaPIP1-6a shares 97%-98% identity at cDNA level and 95-96% at protein level to DQ867075, a TaPIP1 (AQP4). The highest identity of the TaPIP2 and TaPIP1 genes obtained from cDNA and gDNA from this study to the Triticum turgidum confirms that these two likely new genes may belong to AA or BB genome. The predicted amino acid identity shows that the obtained genes may have similar function to the EST (DQ867078 and DQ867075). The alternate splice variant (232 aa) which lacks TMH2, LB and TMH3 is 97% identical to the EST.
### Chapter 2  
Cloning and sequencing of wheat PIP genes

**A**

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Isolated gene</th>
<th>Triticum aestivum</th>
<th>Triticum turgidum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% identity to</td>
<td>% identity to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cDNA (DQ867078)</td>
<td>Protein (ABJ96815)</td>
</tr>
<tr>
<td><strong>PIP2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaPIP2-9a</td>
<td>TaPIP2 cds</td>
<td>95</td>
<td>99</td>
</tr>
<tr>
<td>TaPIP2-9a</td>
<td>TaPIP2 AS</td>
<td>98</td>
<td>97</td>
</tr>
<tr>
<td>TaPIP2-9a</td>
<td>TaPIPA</td>
<td>95</td>
<td>98</td>
</tr>
<tr>
<td>TaPIP2-9b</td>
<td>TaPIPB</td>
<td>97</td>
<td>99</td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Isolated gene</th>
<th>Triticum aestivum</th>
<th>Triticum turgidum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% identity to</td>
<td>% identity to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cDNA (DQ867075)</td>
<td>Protein (ABJ96812)</td>
</tr>
<tr>
<td><strong>PIP1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaPIP1-2a</td>
<td>TaPIP1 cds</td>
<td>97</td>
<td>96</td>
</tr>
<tr>
<td>TaPIP1-6a</td>
<td>TaPIPC</td>
<td>97</td>
<td>95</td>
</tr>
<tr>
<td>TaPIP1-2a</td>
<td>TaPIPD</td>
<td>98</td>
<td>96</td>
</tr>
</tbody>
</table>

**Table 2.11** % identity of genes isolated from gDNA and cDNA to EST
2.3.5 Relationship between the TaPIP2-9a cds and TaPIP1-2a cds is orthologous to Barley and Rice

The alignment of predicted amino acid sequences of TaPIP2-9a cds and PIP1-2a cds with Rice (Oryza sativa), Arabidopsis (A.thaliana), Barley (Hordeum vulgare) and wheat (Triticum aestivum and Triticum turgidum), indicated they are orthologous with barley and rice genes shown in figure 2.12.

The amino acid sequences of the Rice (31 AQPs), Barley (24 AQPs), wheat (29 AQPs) and Arabidopsis aquaporins (35 AQPs) were retrieved from NCBI protein database. Among twenty nine wheat AQPs, twenty seven were from T.aestivum and two (ABW34454 and ABW34453.1) were from T.Turgidum (Ayadi et al., 2011). Twenty two partial putative wheat AQPs are submitted (Forrest and Bhave, 2008) and the rest five full length wheat AQP sequences, AAM00369.1 and AAM00368.1 (Aroca et al., 2005), AAF61465.1, AAF61464.1 and AAF61463.1 submitted by Gao,Y.P. and Gusta,L.V.,2009. The partial wheat AQP sequences were not used for phylogenetic tree construction and therefore nine full length wheat AQPs were used in this study. The six novel wheat AQPs TaAQP1 (PIP1), TaAQP2, 3 and 4 (PIP2), TaAQP5 (TIP1) and TaAQP6 (TIP2) (Zhang et al., 2008) in the accession numbers DQ867075, DQ867076, DQ867077, DQ867078, DQ867079 and DQ867080 were not listed in the protein database, the four PIP sequences (TaAQP1 – 4) were translated to amino acid sequences and used in this study for nucleotide sequence comparison and in the phylogenetic tree construction.

2.3.5.1 The cDNA and alternate splice variant of TaPIP2-9a gene

The putative protein sequence of TaPIP2-9a (cDNA) scored highest protein identity to barley HvPIP2-2 (up to 97% protein identity), to rice OsPIP2-6 (up to 87% protein identity) and Arabidopsis AtPIP2-8 (up to 78%
Chapter 2 Cloning and sequencing of wheat PIP genes

protein identity). The identity values represented here are generated by clustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

The alternate splice variant was speculated to be a likely non functional or less functional because of the lack of TMH2, Loop B and TMH3, where Loop B carries an NPA motif and the absence of TMH2 and TMH3 will fail to form hourglass structure. The study of importance of NPA motifs was investigated by Jiang in 2009 using human aquaporin1 (hAQP1). The deletion of NPA1 or NPA2 motif (single deletion) significantly reduced human AQP1 water permeability to 50% compared to wild-type AQP1 whereas the water permeability was little affected in the double deletion (NPA1 and NPA2) lines (Jiang, 2009). Further the expression of all three mutants indicated identical protein pattern and similar pattern of expression in the plasma membrane compared to wild type AQP1 (Jiang, 2009). Hence the experimental evidence suggests that NPA motifs are essential for water transport but not for the expression of hAQP1. Therefore the alternate splice variant of TaPIP2-9a AS which was obtained from root cDNA would have similar expression pattern to the wild type TaPIP2-9a but the ability of the alternate splice variant to transport water may not be likely, but this splice variant TaPIP2-9a AS could participate in the regulatory function or during dimerization of the aquaporin proteins.

2.3.5.2 The gDNA of TaPIP2-9a gene

The putative protein sequence of TaPIP2-9a and TaPIP2-9b (gDNA) scored highest protein identity to barley HvPIP2-3 (up to 72% and 73% protein identity), to rice OsPIP2-6 (up to 87% and 89% protein identity) and to Arabidopsis AtPIP2-8 (up to 79% and 80% protein identity). The gene structure of TaPIP2-9 (Figure 2.7 (A1) & (A2)) is similar to OsPIP2-7 (intron III) (Forrest and Bhave, 2008).
2.3.5.3 The cDNA TaPIP1-2a gene

The putative protein sequence of TaPIP1-2a (cDNA) scored highest protein identity to barley HvPIP1-1 (up to 98% protein identity), to rice OsPIP1-1 and OsPIP1-2 (up to 89% and 90% protein identity) and Arabidopsis AtPIP1-5 (up to 87% protein identity). The results were confirmed in a bootstrapped phylogenetic tree (Figure 2.12).

2.3.5.4 The gDNA of TaPIP1 gene

The putative protein sequence of TaPIP1-6a and TaPIP1-2a (gDNA) scored highest protein identity to barley HvPIP1-1 (up to 97% and 97% protein identity), to rice OsPIP1-1 and OsPIP1-2 (up to 90% and 91% protein identity) and Arabidopsis AtPIP1-5 (up to 88% protein identity). Further, TaPIP1-6a and TaPIP1-2a is similar in their gene structure (Figure 2.11A and B) to OsPIP1-1 (three introns) (Forrest and Bhave, 2008).
Figure 2.12 The phylogenetic tree (ClustalW alignment of deduced amino acid sequence) demonstrating orthology between rice, barley and *A.thaliana* PIPs.

The full length coding sequences of two wheat aquaporins (*TaPIP1-2* cds and *TaPIP2-9a* cds) and alternate splice variant of *TaPIP2-9a* (*TaPIP2-9a AS*) are highlighted in ash. The length of each branch is proportional to the divergence of that protein sequence from other members of the family. *TaPIP2-9a* cds and *TaPIP1-2a* cds deduced amino acid sequences were obtained from root cDNA in this study. The accession numbers of the protein sequences is in the appendix 4, page no 169.
2.3.6 Prediction of substrate specificity and Mercury sensitivity of wheat AQP proteins

The amino acid sequences predicted from their cDNAs of *TaPIP2-9a* and *TaPIP1-2a* indicated that the PIPs exhibited the characteristic six transmembrane helical domains (TMHs) (Figure 2.13A and B), Loop B and E had highly conserved NPA motifs and therefore the "hourglass structure" (Jung *et al.*, 1994) of MIPs is expected for these proteins.

Figure 2.13 Kyte-doolittle plots of Wheat aquaporins

A and B: A transmembrane proteins, aquaporins contain both hydrophobic and hydrophilic regions. The Kyte-doolittle plot, which predicts the regions of hydrophobicity (window = 19), suggests *TaPIP2-9a* cds (A) and *TaPIP1-2a* cds (B) has six transmembrane.
The phylogenetic analysis of the deduced amino acid of known glycerol and water transport human AQPs with the two wheat AQPs (figure 2.14) clearly indicates that the two genes TaPIP1-2a and TaPIP2-9a are water transport proteins, clustering with hAQP1. Further the conserved amino acid residues distinguishing AQPs and GLP (Glycerol transport proteins) (Refer table 2.1) proteins (figure 2.14) confirm that the TaPIP1-2a and TaPIP2-9a genes are not aquaglyceroporins. The table 2.1 adapted from (Quigley et al., 2001) was used to verify the wheat TaPIP1-2a and TaPIP2-9a genes for substrate specificity.

![Phylogenetic tree](image)

*Figure 2.14 Phylogenetic tree (ClustalW alignment of deduced amino acid sequence) demonstrating the substrate specificity of AQPs.* The proteins boxed with lines are the water transporting AQPs and the aquaglyceroporins are boxed by dotted lines.

The mercury (Hg) sensitive residue corresponding to Cys189 of human AQP1 (Preston and Agre, 1991, Preston et al., 1994) was replaced by Thr in TaPIP1-2a and TaPIP2-9a at position 279 in the figure 2.15. The Cys-189 in hAQP1 is located at 3 residues prior to the second NPA sequence and directly associated with pore formation, and substitution with tryptophan has been shown to block the channel, whereas Cys11 (highlighted in blue) is predicted to be the Hg sensitive site for hAQP3, near the NH$_2$ terminal (Kuwahara et al., 1997). The Hg sensitive site Cys11 in AQP3 might not directly associated with pore formation and its significance remains unclear at present (Kuwahara et al., 1997). AQP7 doesn't
Chapter 2  Cloning and sequencing of wheat PIP genes

contain a Cys residue in this region and was reported to be Hg insensitive (Ishibashi et al., 1997).

Less information is known about the potential metal sensitive residues in plant AQPs; NtAQP1 (Biela et al., 1999), γ-TIP (Daniels et al., 1996) and δ-TIP were found to be Hg insensitive. The putative mercury coordination sites in plant aquaporins are predicted to be in the TMH2 and TMH3 (Guo et al., 2006a), shown in (Appendix 3), with corresponding rice AQPs, in the positions Cys119 and Cys155 for TaPIP2-9a and Cys119, Cys155 and Cys160 for TaPIP1-2a in the alignment. In the putative protein sequences of TaPIP2 the mercury sensitive positions are Cys94 and Cys130, for TaPIP1-2a the mercury sensitive positions are Cys106, Cys142 and Cys147.

2.3.7 Prediction of ammonia transport sites in the wheat AQP proteins

The Obtained TaPIP2-9a and TaPIP1-2a putative amino acid sequence was analyzed for ammonia transport. The known ammonia transporting AQP in humans are hAQP1, 3 and 8; in plants TaTIP2-2 and NtTIPa. The prospective amino acid residues for ammonia facilitating plant aquaporins, TaTIP2-2 were Ile97, Leu168 and Gly224 and the non ammonia facilitating AtPIP1-2 carries Leu 129, Val 209 and Ala268 and NiPIP2-1 carries Val117, Val197 and Ala256 (Muhammad, 2011). The respective positions were analyzed for the TaPIP2, which carries Val117, Val197 and Ala256 and TaPIP1-2a which carries Leu129, Val209 and Ala268 as same as non ammonia facilitating AtPIP1-2 and NiPIP2-1. Therefore the putative TaPIP2-9a and TaPIP1-2a protein are not likely to facilitate ammonia transport.

2.3.8 Prediction of important amino acid residues in wheat aquaporins compared to rice

The residues differentiating PIP1 and PIP2 proteins are shown in appendix 3. The phosphorylation sites are also predicted which are involved in gating (Refer section 1.6.1, Chapter 1) and regulation of aquaporins (Refer section 1.13.1, Chapter 1). The experimental evidence in A.thaliana suggests that there were
multiple phosphorylation sites in the C-terminal of \textit{AtPIP2}, whereas no phosphorylation sites were detected in the \textit{AtPIP1} (Prak et al., 2008). Sequence alignment of rice and full length wheat aquaporins \textit{TaPIP1-2a} and \textit{TaPIP2-9a} (Appendix 3) shows the potential ser residue in the C-terminal of PIP2 that could be phosphorylated. The lack of Ser residue in the C-terminal of \textit{TaPIP1-2a} obtained from this study may also transport less water, like other PIP1 plant aquaporins reported in previous studies (Chaumont et al., 2000).

### 2.3.9 Candidate aquaporins to facilitate water uptake in barley

Sequence comparison between barley and wheat PIPs shows that \textit{TaPIP1-2a} and \textit{TaPIP2-9a} share highest sequence identity with \textit{HvPIP1-1} and \textit{HvPIP2-2}. \textit{HvPIP2-2} display water channel activity (Besse et al., 2011) and showed significant differences in expression between root zones (transition zone, mature zone and lateral root) (Knipfer et al., 2011). \textit{HvPIP2-2} and \textit{HvPIP1-1} were expressed higher in the elongation zone compared to mature and emerged leaf tissue (Besse et al., 2011), where the expression of \textit{HvPIP1-1} accounted for 90% among all \textit{HvPIP1s} (Wei et al., 2007). The high expression levels and heteromerization (Fetter et al., 2004, Zelazny et al., 2007) of the \textit{HvPIP2} proteins may partially compensate the low water transport activity of \textit{HvPIP1-1} (Wei et al., 2007). \textit{HvPIP2-2} was mainly localized to the plasma membrane; \textit{HvPIP2-2} and \textit{HvPIP1-1} are highly expressed in the epidermis (Wei et al., 2007). \textit{TaPIP2-9a} cds and \textit{TaPIP1-2a} cds may exhibit similar function and expression pattern to \textit{HvPIP2-2} and \textit{HvPIP1-1}.  

91
### Chapter 2

Cloning and sequencing of wheat PIP genes

<table>
<thead>
<tr>
<th>200</th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>GlpF, Ecoli</td>
<td>LAGTFSTYPN PHINFVQAFA VEEVVTATLM GLTA TDDG N---GVPVRGP LAPELITLLLI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hAQP3</td>
<td>T.I.A...S G.LDMNG.F DQVF.G .S.I VCV.IV.PY ---NPVPRG TAETV.IV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hAQP7</td>
<td>T.I.A..LP D.MTLWKG.L N.AWL.GN.Q LCLF.I.Q.E ---NPALPG TEEV.IIV.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hAQP9</td>
<td>T.H.I..A..YLSL...D QCV.LG.RS .L IIVF.IF.SR ---IGAPR.G EELT.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hAQP10</td>
<td>T.S.I.A..A.YLSLNNG.L DQVLG.GM.I VGL.II.RR ---KSPVAG.E VVVV.M.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hAQP1</td>
<td>----RNDLA DGV.GS.GLG I.LTTLV.LCV..T.RD---L.G S...A...SV --</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| TaPIP2-9a | ---ANVVA SGFSRGT.LG A.I.1T.V...Y...V...S...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...V...
Figure 2.15 Alignment of glycerol, water transport and Mercury sensitive AQPs.

Alignment and consensus sequence of human aquaporins, bacterial aquaporin and wheat aquaporins. Sequence data used in the figure are referenced in PubMed protein query under Accession Nos are AAA58425.1 (AQP1), NP_004916 (AQP3), NP_001161 (AQP7), BAA24864 (AQP9), CAH70483 (AQP10), bacterial glycerol transporter andYP_859517.1 (glpF). The protein legends highlighted in grey are water transport proteins. P1-P5 (Position 1 to 5) is conserved residues (Quigley et al., 2001) that distinguish functions, differentiated in green (glycerol transporters) and red shading (water transporters) (Quigley et al., 2001). Numberings are arbitrary.
Chapter 2  Cloning and sequencing of wheat PIP genes

2.4 Summary

Two full length cDNA encoding AQPs were isolated from the common wheat cultivar Cranbrook, of ORF lengths of 867 and 897 bp in length. All (TaPIP1-2a cds and TaPIP2-9a cds) of these cDNA sequences are not very different from the known wheat AQP genes in sequence comparison analysis. These genes were temporarily designated as TaPIP1-2a cds and TaPIP2-9a cds. Sequence alignment analysis was conducted for the predicted proteins coded by the two TaPIP1-2a cds and TaPIP2-9a cds genes, together with the known 11 rice PIPs, 10 barley PIPs, 13 A.thaliana PIPs and 26 wheat PIPs (alignment data not shown). The two wheat TaAQPs were classified into PIP1 and PIP2 subgroups. Only nine complete sequences of PIPs from T.aestivum (accession numbers: AF139814 (TaPIP1), AF139815 (TaPIP2), AF139816 (TaPIP3), AF366564 (TaPIP1), AF366565 (TaPIP2), DQ867075 (TaPIP1), DQ867076 (TaPIP2), DQ867077 (TaPIP2) and DQ867078 (TaPIP2) and two from T.Turgidum (accession numbers: EU162655 (TtPIP2) and EU162654 (TtPIP1)) were found, which represents three subgroups of PIPs, PIP1, PIP2 and PIP3 respectively, based on the annotation. In the present study two full length AQP genes (TaPIP1-2a cds and TaPIP2-9a cds) and one alternate splice variant of TaPIP2-9a gene was isolated from root cDNA of early seedling stage of wheat plant. Four full length AQP genes (TaPIP2-9a, TaPIP2-9b and TaPIP1-2a, TaPIP1-6a) from genomic DNA were isolated from T.aestivum cv.Cranbrook and the gene structure was constructed by intron/exon nucleotide sequence comparison.

The expression analysis of putatively expressed rice AQP genes revealed that most of the AQPs are expressed at the seedling stages and therefore wheat AQPs may also have the same expression pattern as rice AQPs. This study also confirmed the expression of a wheat PIP2 gene at early seedling stage and not at 21-day-old cDNA. Computational prediction revealed that the two TaAQPs containing six conserved transmembrane regions (fig 2.5) and the residues that differentiates PIP1 and PIP2 in plant AQPs are conserved with the Rice MIPs (alignment file in appendix 3, Page
The putative protein sequences of TaPIP2-9a (cDNA), TaPIP2-9a (gDNA) and TaPIP2-9b (gDNA) scored highest identity with 98% and TaPIP1-2a (cDNA), TaPIP1-2a (gDNA) and TaPIP1-6a (gDNA) scored highest identity of up to 92% with wheat Triticum turgidum TtAQP2-1 and TtAQP1-1. Therefore the amplified TaPIP1 and TaPIP2 gene may be located in the AABB genome of the common wheat (Triticum aestivum). The putative protein sequences of TaPIP1 and TaPIP2 gene shared highest identity of up to 97% with barley HvPIP1-1 and HvPIP2-2 (Hordeum vulgare) and thus the wheat aquaporin genes obtained from this study are orthologous to these barley genes. The sequence data yielded in this work on full length wheat aquaporin genes TaPIP2-9a cds, TaPIP1-2a cds and likely new homeologs of TaPIP1-2, TaPIP1-6 and TaPIP2-9 (TaPIP2-9a and TaPIP2-9b) provides scope for the heterologous expression of these proteins to test the function of the protein. In addition the above mentioned sequence data can also be used to identify the chromosomal locations by physical mapping using nullisomic tetrasomic lines.

The comparison of the predicted primary amino acid sequence analysis with the known human aquaporins also revealed that the two full length wheat aquaporin obtained from this study likely transport water and not glycerol. Mercury sensitive sites were also predicted using rice and human aquaporin genes. The presence of mercury sensitive cysteines in the wheat AQPs corresponding to the rice AQPs suggests that these two TaPIP1 and TaPIP2 may represent mercury sensitive AQPs. Further these two genes was also analysed for residues predictive for the ability to transport ammonia by comparing the amino acid sequences of the wheat AQPs with the known plant ammonia transporters (AtPIP1-2 and NtPIP2-1), the result showing that the wheat AQPs TaPIP1-2a and TaPIP2-9a could possibly be a non ammonia transporter.
3

Heterologous expression of aquaporin genes from human and wheat in Baculovirus infected insect cells.
Chapter 3  
Heterologous expression of AQP genes from human and wheat in Baculovirus infected insect cells.

3.1 Introduction

The Baculovirus expression system was chosen as the method for expression of the aquaporins in insect cells as mammalian (human) aquaporins have been successfully expressed in this system (Yakata et al., 2011, Yang et al., 1997, Werten et al., 2001). Baculoviruses have a large, double-stranded, covalently-closed, circular DNA genome of between 88 and 200kbp (Arif, 1986) and have been isolated from invertebrates, mostly insects. Baculovirus infection has been reported in over 600 species of insects (Martignoni and Iwai, 1986). Baculovirus was developed as an expression vector by Summers and Smith, 1979. Baculoviruses can express genes from bacteria, viruses, plants and mammals at levels from 1-500 mg/liter, although making predictions is difficult in the case of membrane proteins. Baculoviruses are considered one of the safest expression systems at the present time to overproduce the protein of interest (McCarroll and King, 1997, Possee, 1997).

The baculovirus system has been extensively used to express functional membrane transporters and channels, including the cardiac Na\(^+\)-Ca\(^+\) exchanger (Li et al., 1993), NaPi-cotransport (Fucentese et al., 1995), Na\(^+\)/glucose transporter (Smith et al., 1992). Human AQP2 (Werten et al., 2001) and AQP4 (Yang et al., 1997) have been expressed in baculovirus expression system.

BacMAGIC DNA (Novagen) was used to generate recombinant baculoviruses for target protein expression in insect cells without the need for time-consuming plaque assays. Bacterial artificial chromosome (BAC) (Figure 3.1) is present in the BacMagic DNA in the place of the polyhedron coding sequence. A compatible transfer plasmid (e.g. pBacPAK9 and pBacPAK His1) containing the target coding sequence is cotransfected with BacMagic DNA into insect cells. In the insect cells, homologous
recombination occurs and the target coding sequence replaces the BAC sequence. Thus recombinant baculovirus is constructed to produce homogeneous population of recombinants.

3.1.1 Phase separation purification of integral membrane proteins

Triton X-114 a non ionic detergent was used to enrich integral membrane protein by phase separation (Bordier, 1981). The integral membrane protein directly interacts with the hydrophobic core of the lipid bilayer. The nonionic detergent solubulises and replaces most of the lipid molecules in the bilayers and forms protein-detergent mixed micelle (Reynolds and Tanford, 1976, Helenius and Simons, 1972). The non-ionic detergents are less likely than the ionic detergents to cause alterations in the state of aggregation or other conformational properties of the protein (Helenius and Simons, 1972, Makino et al., 1973). Many other integral membrane proteins have been separated by phase partitioning using Triton X-114 including: Synaptobrevin, an integral membrane protein of 1800Dalton in small synaptic vesicles of rat brain (Baumert et al., 1989), integral membrane protein located in the apical complex of *Plasmodium falciparum* (Peterson et al., 1989) and in virulent Treponema pallidum (Radolf et al., 1988).

3.1.2 La Protein

La proteins are nuclear phosphoprotein, ubiquitous in eukaryotic cells and are associated with the 3’termii of many newly synthesized small RNAs. La protein is located in the nucleus, within the nucleus it is both nucleoplasmic and nucleolar; and La protein is also located in cytoplasm (Wolin and Cedervall, 2002). Histidine tagged La protein was used as a negative control to test the efficiency of the phase partitioning method and used as a positive control for the IMAC –pull down and western blot. La protein is not an integral membrane protein and therefore La protein will act as a suitable control for this study to test the efficiency of the method.
Chapter 3  
Heterologus expression of AQP genes

The specific aims of this section of the study were to construct histidine tagged and untagged wheat aquaporins *TaPIP2-9a, TaPIP1-2a* and mammalian aquaporin hAQP3 transfer vectors using the pBacPAK His1 and pBacPAK 9 plasmids. These transfer vectors were to be used to construct recombinant baculoviruses. This entails verification of the gene of interest in the baculoviral genome by viral DNA extraction and PCR using Bac primers and sequence specific primers; and immunoblotting to verify the expression of Histidine tagged wheat and mammalian aquaporins using Monoclonal Anti-Ploy histidine Antibodies.

3.2 Methods

3.2.1 Sf21 cell culture

Sf-21 insect cells (Invitrogen) were cultured in SF-900II serum free media (SFM) (Invitrogen) in 100 mL suspension cultures in 1 L shaker flasks, and incubated at 28°C with shaking at ~110 rpm. Cells were routinely subcultured at a density of 0.5×10^6 cells/mL every 3-4 days when the cell density reached ~4×10^6 cells/mL.

3.2.2 Co-transfection

Genejuice Transfection Reagent (Novagen) enables highly efficient DNA transfer in the insect cells to facilitate homologous recombination (Figure 3.1). The method was modified from the BacMAGIC (Novagen) method by halving the volumes of reagents. For each reaction 95 µL of sterile water, 5 µL of insect gene juice (transfection reagent, Novagen), 2.5 µL of BacMAGIC DNA (50ng) and approximately 500 ng of sterile plasmid transfer vector (10 µL of ethanol precipitated plasmid sterile DNA) was added into a sterile 1.5ml microcentrifuge tube to the total volume of 112.5 µL. Co-transfection reaction mixture (GeneJuice reagent/DNA mixture) was mixed gently by pipetting and the mixture was incubated at RT for 20 minutes. The GeneJuice reagent /DNA mixture was mixed into a sterile 1.5ml microcentrifuge tube to the total volume of 112.5 µL. Co-transfection reaction mixture (GeneJuice reagent/DNA mixture) was mixed gently by pipetting and the mixture was incubated at RT for 20 minutes. The GeneJuice reagent /DNA mixture was then transferred into a 25 cm² T-flask previously seeded at the density of 1×10^6 Sf-21 cells in 5 ml SF-900II SFM medium (Invitrogen). The co-transfection was incubated for 5-7 days at 28°C. The infected insect cells were harvested by centrifugation (2000 rpm for 10
Chapter 3  

Heterologous expression of AQP genes

mins), the supernatant (seed stock) was stored at 4°C and pellet was stored at -20°C and used for viral DNA extraction.

![Diagram of recombinant virus construction]

**Figure 3.1 Construction of baculovirus recombinants with BacMagic systems by homologous recombination.** (Image adapted from Novagen, user protocol TB459 Rev.E 1208)

### 3.2.3 PCR verification of recombinant virus

#### 3.2.3.1 Viral DNA extraction by phenol chloroform method

Baculoviral DNA of wheat aquaporin was extracted using a phenol chloroform method. The cell pellet obtained from 5 mL of co-transfected sample (section 3.2.2) was used to extract DNA. The frozen cell pellet was resuspended completely with 250 µL of Cell resuspension from Wizard Genomic DNA Purification Kit (Promega) solution and vortexed briefly. 200 µL of Cell Lysis buffer from Wizard Genomic DNA Purification Kit (Promega) was added and mixed by inverting the tube several times and an equal volume of Phenol chloroform was added and vortexed for 10 sec and spun at 12,000 rpm 4°C for 5 min. The aqueous phase containing the DNA was carefully removed and transferred to a new 1.5 mL eppendorf tube. DNA was back extracted by adding 260 µL of 1X TE buffer to the bottom phase and spun at 12,000 rpm 4°C for 5 min. The aqueous phase containing the DNA was carefully removed and pooled with the first aqueous was obtained from
Chapter 3 Heterologus expression of AQP genes

the first extraction. The obtained DNA was ethanol precipitated. To the DNA sample 1/10th volume 3M sodium acetate (pH 5.5) was added and vortexed briefly or by flicking the tube several times with finger. 3 volumes of ice cold 100% AR grade ethanol was added and mixed by vortexing and precipitated at -20ºC for 30 min. The DNA precipitate was spun at 14,000 rpm for 15 minutes at 4 ºC and the supernatant carefully aspirated using a micropipette or Pasteur pipette. At least 1 volume of 70% ethanol was added and the tubes inverted several times. The tubes were spun at 14,000 rpm for 15 minutes at 4ºC; carefully aspirated the supernatant and the obtained pellet was air dried and resuspended in 20 µL of sterile water. The obtained DNA was verified by running DNA gel electrophoresis.

3.2.3.2 Viral DNA extraction by isopropanol method

Baculoviral DNA of mammalian aquaporin was extracted using both phenol chloroform and isopropanol method. Refer section 3.2.3.1 for the protocol until back extraction and few modifications in the last step. 400 µL of the pooled supernatant was transferred to the new microcentrifuge tube and 600 µL of isopropanol and mixed by inverting the tube and centrifuged at 14,500 rpm for 3 min at RT. The supernatant was discarded without disturbing the pellet and the pellet was washed with 100 µL of 70% ethanol. Ethanol was removed carefully and the pellet was air dried. Pellet was resuspended in 100 µL of sterile water and incubated at 65°C for 1 hour or at 4°C overnight and store the sample at -20 ºC.

3.2.3.3 PCR verification of recombinant baculoviruses

PCR was performed from the DNA in the obtained by Phenol chloroform and isopropanol method using BacPAK vector based primers Bac1 and Bac2 and recombinant aquaporin specific primers TaPIP2F and TaPIP2R1 for PIP2 viral DNA; TaPIP1F and TaPIP1R for PIP1 viral DNA and AQP3F and AQP3R for hAQP3 viral DNA, for primer sequence see (Table 4.1). PCR reaction mixture was prepared in the total volume of 25 µL, includes 12.5 µL of Mango mix (Bioline); 1 µL of 10mM Bac1 and Bac2
primers; 1 µL of 100-150 ng of template viral DNA and 10.5 µL of sterile water. A 'uninfected insect cell' negative control was included in each PCR run. PCR condition was Step 1: Initial denaturation at 95°C for 5 min; Step 2: 30 cycles of repeated denaturation at 95°C for 40 sec, annealing at 55°C for 40 sec and extension at 72°C for 40 sec. A final extension at 72°C for 10 minutes was applied. The PCR products were then verified by 1% agarose gel electrophoresis (see section 2.2.7.2.2).

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ – 3’)</th>
<th>Direction</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP2 Fab</td>
<td>AAGaattcATGGCCAAGGAGGTGAGCGAG</td>
<td>Forward</td>
<td>Verify the wheat AQP construct, TaPIP2-9a by PCR</td>
</tr>
<tr>
<td>TaPIP2 Ra2</td>
<td>AAGaattcTTAATTATTGCGCCGGGGG</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>TaPIP1 F</td>
<td>AAGaattcATGGAGGGCAAGGAGGAGGAC</td>
<td>Forward</td>
<td>Verify the wheat AQP construct, TaPIP1-2a by PCR</td>
</tr>
<tr>
<td>TaPIP1 R</td>
<td>AAGaattcTTAGGACTGGTTCTGAACGGGATCGC</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hAQP3 F</td>
<td>ATGGGTCACTGCAAGGAGGC</td>
<td>Forward</td>
<td>Verify the human AQP3 construct by PCR</td>
</tr>
<tr>
<td>hAQP3 R</td>
<td>TCAGATCTGCTCCTTGCTC</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>Bac 1</td>
<td>AACCATCTCGCAAATAAATA</td>
<td>Forward</td>
<td>Sequencing the Baculoviral constructs</td>
</tr>
<tr>
<td>Bac 2</td>
<td>ACGCACTAATCTAGCGCTT</td>
<td>Reverse</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1 Primers used to verify the Viral DNA insert

3.2.4 Expansion of recombinant baculoviruses

The verified recombinant baculoviruses were expanded in 100 mL sf-21 suspension cultures with 0.5×10⁶ cells/mL in SF-900II media (Invitrogen). Cells were infected with 100 µL of co-transfection supernatant (therefore at low M.O.I) and incubated at 28°C for 4-5 days with shaking at 120 rpm. The infected Sf-21 cells were collected by centrifugation (2000 rpm for 10 mins at 4°C). The obtained viral supernatant was stored at 4°C wrapped with aluminum foil and the cell pellet was discarded.
3.2.5 Infection of recombinant baculoviruses

The recombinant baculoviruses were used to infect 200 mL sf-21 cultures with $2.0 \times 10^6$ cells/mL in Sf-900II media (Invitrogen). Cells were infected with 5 mL of viral inoculum (MOI of approximately 1) obtained from expansion and incubated at 28°C for 3 days with shaking at 120 rpm. The infected Sf-21 cells were collected by centrifugation (2000 rpm for 10 mins at 4°C). The infected cell pellet was further used for protein extraction.

3.2.6 Membrane protein extraction

The infection cell pellet obtained was resuspended in 10 mL of ice cold 50 mM phosphate buffer (pH 7) and incubated for 30 min at 4°C. The incubated mixture was spun at 2000 rpm for 5 min, 4°C. The supernatant was discarded and the obtained pellet was resuspended in 10 mL ice cold 50 mM phosphate buffer (pH7) and mixed thoroughly by pipetting.

3.2.7 Homogenization using a syringe and needle

The final suspension obtained from (Section 3.2.6) was subjected to homogenization by passing the lysate through a 25-gauge (0.5 mm) needle, attached to a 10 mL sterile plastic syringe, at least 10-12 times or until a homogeneous lysate was achieved.

3.2.8 Sorvall centrifugation

The homogeneous lysate obtained as a result of homogenization was then centrifuged at 4000 rpm for 5 min. The supernatant was transferred into Sorvall centrifuge tubes and spun at 20,000 rpm, for 20 minutes, 0°C in the Sorvall (SS34 rotor) centrifuge (Thermo Fisher Scientific Inc.). The supernatant (clear) was discarded and pellet was resuspended in 400 µL of cold phase separation buffer (10 mM Tris HCl (pH 7.4), 150 mM NaCl, 1% Triton X-114) and aliquoted into 200 µL, 100 µL and 100 µL in three separate sterile 1.5 mL microcentrifuge tubes for phase separation, Bradford assay and SDS PAGE.

3.2.9 Bradford assay

The amount of total protein in the membrane preparations was estimated using a micro titer tray-based Bradford assay. BSA standards were prepared in the concentration ranging from 0 to 5 mg/mL. The protein
samples (crude extract) were diluted in 1 in 10 dilutions (10 µL in 100 µL) with sterile water and equal volume (100 µL) of Bradford reagent (recipe) was added. The standard protein and the test protein samples were loaded onto the 96 well microtitre plate (Invitrogen) and absorbance was read at 591 nm.

### 3.2.10 Phase Separation using detergent Triton X-114

Triton X-114 is a nonionic detergent, homogeneous at 0°C but separates into aqueous phase and detergent phase above 20°C (Bordier, 1981). This method is adopted from the published article of Bordier, 1981. The protein sample was then separated on a cushion of 6% sucrose (6% sucrose, 10mM Tris HCl, pH 7.4, 150 mM NaCl, 0.06% Triton X-114). 300 µL of 6% sucrose cushion was placed in the 1.5 mL sterile microcentrifuge tubes. The 200 µL (obtained as a result of Sorvall centrifugation) of clear protein sample was then overlaid on this sucrose cushion, and the tube was incubated for 3 min at 30°C. Clouding of the solution occurred. The tube was centrifuged at 2500 rpm for 3 min at room temperature. The detergent phase (pellet, can be seen as an oily droplet) settled at the bottom of the tube and the aqueous phase (supernatant) on top. The obtained aqueous phase was then removed carefully and overlaid on to the 300 µL of fresh sucrose cushion in a new 1.5 mL microcentrifuge tube and was incubated for 3 min at 30°C for phase separation to occur and the detergent phase (first fraction) was stored at 4°C. The incubated reaction mixture was transferred on to the detergent phase (first fraction) which was stored at 4°C and centrifuged at 3000 rpm for 3 min. The resulting aqueous phase (supernatant) was transferred on to a fresh 1.5 mL tube and the detergent phase (second) (pellet) was split into two tubes, 20 µL each. One of the tubes containing 20 µL of the detergent phase (second) was diluted with 20 µL of phase separation buffer with 1% Triton X-114 and the other tube with 20 µL of the detergent phase was undiluted. The diluted and undiluted fractions of detergent phase (second) were treated with IMAC resins for the purification of the His-tagged protein.

The aqueous phase (second) was treated with 200 µL of phase separation buffer containing 2% Triton X-114 and centrifuged at 3000rpm for 3min at room temperature; the resulting detergent phase (third) was discarded and the aqueous phase (third) was split into two tubes, 20 µL in each tube. To one of
Chapter 3  
Heterologus expression of AQP genes

the tubes containing 20 µL of the aqueous phase (third) was diluted with 20 µL of phase separation buffer with 1% Triton X-114 and the other tube with 20 µL of the aqueous phase (third) was undiluted. Aliquots of separated phases were analyzed by SDS PAGE electrophoresis.

3.2.11 Affinity pull-down using IMAC (Immobilized Metal ion Affinity Chromatography)

IMAC is widely used for purifying proteins that have affinity for metal ions, such as histidine tagged proteins. IMAC is based on the specific interaction between transition metal ions (Cu$^{2+}$, Zn$^{2+}$, Ni$^{2+}$ or Co$^{2+}$). The Ni-NTA (Nickel-Nitrilotriacetic Acid) affinity purification method was used to purify the desired protein. 20 µL of nickel sepharose beads (GE Healthcare) was placed in a sterile 1.5 mL microcentrifuge tube and was centrifuged at 1000 rpm for 1 min at RT. The supernatant was aspirated carefully without disturbing the beads and the beads was resuspended in 25 µL of wash buffer (Phosphate Buffered Saline, PBS) and spun at 1000 rpm for 1 min. This step was repeated two times and protein sample (undiluted and diluted detergent phase (second)) was then added on to the washed nickel beads and incubated at the RT for 20 mins with end on end rotational wheel (Chiltern Scientific). The incubated samples were centrifuged at 2000 rpm for 5min, the supernatant was discarded and 25 µL of phase separation buffer with 1% Triton X-114 was added to the nickel beads and centrifuged at 1000 rpm for 1 min at RT. The supernatant was discarded and washed twice with 50 mM phosphate buffer at 1000 rpm for 1 min. The supernatant was discarded after wash and the protein was run on the SDS-PAGE by adding 25 µL of 2X Sample Reducing Buffer (100 mM Tris-Cl (pH 6.8), 4% (w/v) SDS, 0.2% (w/v) BPB, 20% (v/v) glycerol, 200 mM mercaptoethanol), heated at 95°C for 3 min and spun at 2000 rpm for 2 min and the supernatant was loaded on to the SDS PAGE.

3.2.12 SDS PAGE

Crude protein samples and affinity purified protein samples were loaded onto 12% SDS-PAGE (with 5% stacking gel and 12% resolving gel) Miniprotein gels (Bio-Rad) and electrophoresed at 150 V for 60 mins in Tris-
Chapter 3  
Heterologous expression of AQP genes

Glycine electrophoresis buffer (25 mM Tris, 250 mM glycine (pH 8.3), 0.1% (w/v) SDS). Fermentas prestained protein ladder PageRuler™ or ColorPlus Prestained Protein Ladder, Broad Range (10-230 kDa) was used as a reference marker. Gels were stained with instant blue (Stratech Scientific).

3.2.13 Analysis of the expression levels of Aquaporins by Western blotting

Samples were loaded on 12% SDS-PAGE (with 5% stacking gel and 12% resolving gel) and electrophoresed at 150 V for 60 mins in Tris-Glycine electrophoresis buffer (25 mM Tris, 250 mM glycine (pH 8.3), 0.1% (w/v) SDS). PageRuler™ prestained protein ladder (Thermo scientific) was used as a reference marker.

Western blotting was used to verify the expression of 6XHis tagged hAQP3 and wheat aquaporins in the baculovirus infected Sf-21 cells. After electrophoresis, the protein bands were transferred to nitrocellulose membranes (Amersham Biosciences Hybond™-ECL™, cat no. RPN203D) using Towbin transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol, pH 8.3). In a semi-dry transfer, the gel and the membrane are sandwiched between two stacks of pre-soaked filter paper in direct contact with plate electrodes (Bjerrum and Schafer-Nielsen, 1986, Kyhse-Andersen, 1984, Tovey and Baldo, 1987). Six filter paper (Bio-Rad), unstained SDS gel and nitrocellulose membrane were equilibrated by pre-wetting in 10mL of Towbin transfer buffer for 15min. Three pre-soaked filter paper was placed on the platinum anode plate of the trans-blot apparatus. Air bubbles in between the filter paper were removed by rolling a test tube on top of the filter paper. Equilibrated membrane was placed on to the thick filter paper and rolled out to remove air bubbles. Equilibrated SDS gel was carefully transferred on top of the membrane and rolled out to remove air bubbles. Three pre-soaked filter paper was placed on top of the gel and air bubbles were removed. Transfer was conducted in semi-dry blotting apparatus (BIO-RAD, Trans-Blot® SD Semi-Dry Transfer Cell) at 15V (BIO-RAD, MODEL 200/2.0 POWER SUPPLY) for 20 mins. The membrane was blocked with 5% non fat dried milk (Skim Milk Powder, Diploma instant) in 10 mL PBS-T (PBS, pH 7.4
Chapter 3 Heterologous expression of AQP genes

with 0.1% Tween 20) for 1hr under shaking condition at room temperature. After three 5 minute washes in PBS-T, the membrane was probed with the antibody specific for Histidine tag: Anti-polyhistidine monoclonal antibody (Sigma Aldrich - H 1029) at a 1:1000 dilution in 10 mL PBS-T containing 1%BSA (Sigma) for 1 hour at room temperature under shaking condition. The membrane was then washed in PBS-T for 15min and the membrane was washed again for 5 minute three times. The membrane was incubated with peroxidase labelled anti-mouse (Sigma) at a 1:1000 dilution in PBS-T containing 1% BSA for 1 hour at room temperature under shaking condition. The membrane was washed four times with PBS-T, first wash for 15 min and subsequent washes for 5 minutes every time with PBS-T. After washing with PBS-T, detection was carried out using 3,3′-Diaminobenzidine tablets (DAB, Sigma Aldrich and Dabco).

3.2.14 **Detection of western blot with DAB**

1 tablet of DAB (10mg) was dissolved in 10 mL of 50 mM phosphate buffer (pH 7) with 10 µL of Hydrogen peroxide (H₂O₂) added immediately before use. The membrane was then incubated with DAB solution for 10 - 20min at room temperature. Bands of red-brown stain were seen from the action of horse radish peroxidase. Aliquots of the separated phases were analyzed together with the starting homogenates by sodium dodecyl sulfate-polyacrylamide gel electrophoresis on slab gels.

3.3 **Result and Discussion**

3.3.1 **Preparation of pBacPAK9 and pBacPAK His1 plasmid DNA**

In this study baculovirus infected insect cells were used to produce recombinant aquaporins. The baculoviral vector pBacPAK9 (5538 bp) and pBacPAK His1 (5414 bp) (Clontech) were used to generate a recombinant expression bacmid. The plasmid DNA was prepared for both untagged pBacPAK9 and tagged pBacPAK His1.
Chapter 3 Heterologous expression of AQP genes

3.3.2 Directional cloning of AQPs into Histidine tagged and untagged baculoviral expression vectors

3.3.2.1 Directional cloning of wheat aquaporins into tagged and untagged baculoviral expression vectors

The full length coding sequence of wheat aquaporins, TaPIP2-9a and TaPIP1-2a cloned in pGEM®-T Easy Vector (Refer Chapter 2, Section 2.2.10) were subcloned in the Baculoviral vectors. BamH1 and EcoR1 were selected as cloning sites in baculoviral vectors pBacPAK9 and pBacPAKHis1to subclone wheat aquaporins TaPIP2-9a and TaPIP1-2a. The forward primers used to amplify TaPIP2-9a and TaPIP1-2a gene from wheat cDNA incorporated a BamH1 restriction site (underlined) (table 4.1) at the 5’ end to match the pBacPAK vectors. The BamH1 restriction site is located one base pair upstream (5’) and the EcoR1 restriction site is located 5 base pairs downstream of the wheat AQP sequences. The Histidine tagged putative amino acid sequence of TaPIP2-9a and TaPIP1-2a, with the predicted Molecular weight of 32 kDa (300 aa) and 32.5 kDa (304 aa), see figure 3.2.

A
MGHHHHHHVVDKLGSMAKEVEEPEHAAPAHKYSDPPAPLPFDMDGELRMRSFY
WALIAEFVATLLFLYITVATVIGYKVRSAAADPCGGVVCVLGIAWAFGGMIFVLVYCTAG
ISGGHINPAVTFTGLLARKVSLRAMLAVMYIVACGIGIGAGIGVGGIGKIMKDAYQANGGG
ANMVSAGFSRGTAFLGVEIVGTFVLVTVFSATDPKRSARDHVPVLAPLPIGFAVFM
VHLATIPITGTGIPARSLGAAVYKNAWDDNHWIFWVGPLVGLAAAAAYHQYIL
RAAAIKALGFSRSSRSN

B
MGHHHHHHVVDKLGSMEGKEEDVRLGANIKSERQPIGTAQQGSEDKDKEEAPPAP
LFEPGELKSWSFYRAGIAEFMATLLFLYVTILTVMGYSAGASKCATVGIGIWAFFSFGG
MIFALVYCTAGISGGHINPAVTFTGLLARKLSLRAVFYIIMQCLGAICGAVVKGFO
QGLYMGNGGANGVAPGYTKGSGLGAELIGTFVLVTVFSATDAKRNARDHVPIL
APLPIGFAVFLVHLGTIPITGTGIPARSLGAAIYNREHAWSDHWFVGVGGPFLAALA
AVYHQVVIRAIPIFFTKS

Figure 3.2 Putative amino acid sequences of Histidine tagged wheat aquaporins.
Chapter 3  

A: The putative protein sequence of TaPIP2-9a is bolded, NPA motifs are highlighted and the Histidine tag of pBacPAK His1 vector is unbolded and underlined.

B: The putative protein sequence of TaPIP1-2a is bolded, NPA motifs are highlighted and the Histidine tag of pBacPAK His1 vector is unbolded and underlined.

The coding regions of TaPIP2-9a and TaPIP1-2a were amplified using gene specific primers (Table 4.1). The forward primer of both wheat aquaporins, TaPIP2-9a and TaPIP1-2a contain BamH1 site. The EcoR1 restriction site flanks the insert at the either side in the pGEM®-T Easy Vector. When digested with BamH1 and EcoR1, pGEM- TaPIP1-2a was expected to give 3015bp and 867bp and pGEM- TaPIP2-9a was expected to give 3015bp and 887bp. These 867 bp and 887 bp inserts were subcloned into a similarly digested pBacPAK9 and pBacPAKHis1 vector to construct tagged and untagged plasmid vector for functional analysis.

![Figure 3.3 Plasmid DNA analysis of the pBacPAK9 constructs of native wheat aquaporins by restriction digestion.](image)

Gel electrophoresis showing (A) Restriction digested pBacPAK9-TaPIP2-9a plasmid DNA. Lane M: GeneRuler™ DNA Ladder Mix; 1 and 2: Plasmid DNA restriction digested with BamH1 and EcoR1, giving fragments of predicted size 5532 bp and 897 bp. 

(B) Restriction digested pBacPAK9-TaPIP1-2a plasmid DNA. Lane M: GeneRuler™ DNA Ladder Mix; 1 and 2: Plasmid DNA restriction digested with BamH1 and EcoR1, giving fragments of predicted size 5535 bp and 864 bp.
3.3.2.2 Directional cloning of mammalian AQP3 into His-tagged and untagged baculoviral expression vectors

Restriction sites compatible with pBacPAK9 and pBacPAKHis1, BamH1 and Xba1, were selected to digest pBacPAKHis1 mammalian hAQP3 cloned in pT-Adv vector. The BamH1 restriction site is located 35 base pairs
Chapter 3  Heterologous expression of AQP genes

upstream (5') and EcoR1 restriction site is located 5 base pairs downstream (3') of the AQP3 sequence. The Histidine tagged amino acid sequence of hAQP3, 300 amino acids with the predicted Molecular weight of 35 kDa, see figure 3.3. The nucleotide sequences of mammalian aquaporin cloned in pT-Adv Vector (Accession no. NM_004925, obtained from Dr.Arun National University of Singapore).

MGHHHHHHHVVDKLGSTSNGRQCGIMGRQKELVSRCEMLHIRYRLLRQALAECFLgL
LVMFGCGLSVAQVVLSRTHGGLINLAFGFAVTGILAGQVSGANLNPATFAMCFLARE
PWIKLPYTQLGLAGVGFLYYDAIWHDNQLFVSGPNATGIFATYPGSHDLMIN
GFFDQIGTASLIVCVAIVDPYNNPVPRLLEAFTVGLVVLVIGTSMSMFNNSGAYVPNAPARELWP
PAFFTALAGWGSAVFTTGQHWWWVPIVSSPLLSIAGVVFVYQLMIGCHLEQPPSNEEENVKL
LAHVKHKEQI

Figure 3.5 Putative amino acid sequence of Histidine tagged hAQP3. The putative protein sequence of hAQP3 is bolded, NPA motifs are highlighted and the Histidine tag of pBacPAK His1 vector is unbolded and underlined.

Recombinant pBacPAK9-hAQP3 was constructed by a two-step process. Mammalian aquaporin ORF was isolated from pT-Adv Vector (kindly supplied by Dr.Arun, National University of Singapore, Singapore) using restriction enzymes BamH1 and Xba1. The same enzymes were then used to digest pBacPAK9 vector (Clontech), a plasmid that contains a Polyhedrin promoter driving the expression of the foreign gene. When digested with BamH1 and Xba1, pT-Adv hAQP3 was expected to give 3.9 Kb and 882. The digested insert, 882bp was then crystal violet gel purified to avoid any mutation (gel picture not shown) and was ligated into the similarly digested and crystal violet gel (data not shown) purified pBacPAK9 and pBacPAKHis1 vector.
3.3.3 Verification of the Recombinant baculovirus

The viral cell pellet obtained after cotransfection (see section 3.2.2) of the viral constructs (pBacPAKHis1-TaPIP2-9a and pBacPAKHis1-TaPIP1-2a) and untagged (pBacPAK9-TaPIP2-9a and pBacPAK9-TaPIP1-2a) were used to isolate viral DNA using phenol chloroform method or isopropanol (section 3.2.3.1) and PCR was performed using expression vector based Bac primers and sequence specific primers (Table 3.1).

3.3.3.1 PCR on Viral DNA

The wheat AQP infected viral DNA was prepared by phenol chloroform method and ethanol precipitated to concentrate the DNA. The mammalian AQP3 infected viral DNA was prepared by phenol chloroform and isopropanol method. The uninfected insect cell was used as a negative control. The concentration of DNA obtained from phenol chloroform method ranging between (160- 200 ng/µL) and isopropanol method (80 - 100 ng/µL). In order to confirm the construction of this baculovirus, viral DNA was isolated from the cell pellet obtained after co-transfection (section 3.2.2).
3.3.3.2 PCR using Baculoviral Vector based primers

The recombinant baculoviruses were verified by PCR on viral DNA using Bac1 and Bac2 primers. These primers bind to the baculoviral DNA regions flanking the inserted AQP genes. PCR was also carried out on uninfected insect cell as a negative control. The Figure 3.7 and 3.8, A and B shows that the lanes 2 and 3 contains a PCR product at approximately 1.1 Kb. This correspond to the expected PCR product of 1088 bp and 1087 bp for untagged and tagged TaPIP2-9a gene in lanes 2 and 3 (Fig 3.7, A), 1058 bp and 1057 bp for untagged and tagged TaPIP1-2a gene in lane 2 and 3 (Fig 3.7, B), 1152 bp and 1182 bp untagged and tagged hAQP3 gene (Fig 3.8, A & B). It was therefore assumed that the wheat AQP genes had been successfully transferred to the baculoviral genome in these six viral isolates.

3.3.3.3 PCR using Sequence specific primers

In order to confirm the above results, a second PCR was carried out with the viral DNA using sequence specific primers TaPIP2-9a and TaPIP1-2a. These sequence specific primers binds within the wheat AQP and this primer pair should result in a PCR product of 897 bp for TaPIP2-9a (Figure 3.7, A, lanes 4 and 5) and 864 bp for TaPIP1-2a (Figure 3.7, B, lanes 5 and 6) and 882 bp for hAQP3. This result provided further confirmation of the wheat aquaporin and mammalian aquaporin gene were successfully subcloned and present in the recombinant baculoviruses. Therefore the verified wheat and mammalian AQP3 viruses were used to infect the insect cells as detailed in section 3.2.5.
Figure 3.7 PCR on Viral DNA to Screen for the His tagged TaPIP2-9a and pBacPAK His1-TaPIP1-2a) and untagged (pBacPAK9-TaPIP2-9a and pBacPAK9-TaPIP1-2a) wheat aquaporin gene cassette.

(A) PCR products of purified viral DNA obtained from tagged and untagged TaPIP2-9a aquaporin constructs. Lanes 1 and 6 are PCR product obtained from uninfected insect cell (negative control); 2 and 3 are untagged and tagged TaPIP2-9a viral DNA using Bac primers; 4 and 5 are untagged and tagged TaPIP2-9a viral DNA using sequence specific primers; M: 1Kb DNA Ladder Mix.

(B) PCR products of purified viral DNA obtained from tagged and untagged TaPIP1-2a aquaporin constructs. Lanes 1 and 4 are PCR product obtained from uninfected insect cell (negative control); 2 and 3 are untagged and tagged TaPIP1-2a viral DNA using Bac primers; lanes 5 and 6 are untagged and tagged TaPIP1-2a viral DNA using sequence specific primers; M: 1Kb DNA Ladder Mix. Note: -ve (negative).
Chapter 3  

Heterologous expression of AQP genes

Figure 3.8 PCR on Viral DNA to Screen for the His tagged (pBacPAKHis1-hAQP3) and untagged (pBacPAK9-hAQP3) mammalian aquaporin Gene Cassette using Phenol Chloroform method (A) and Isopropanol method (B).

(A) PCR products of purified viral DNA obtained from tagged and untagged hAQP3 constructs by Phenol chloroform method. Lanes 1 and 6 are PCR product obtained from uninfected insect cell (negative control); 2 and 3 are untagged and tagged hAQP3 viral DNA using Bac primers (1037 bp and 1087 bp); 4 and 5 are untagged and tagged hAQP3 viral DNA using sequence specific primers (882bp); M: 1Kb DNA Ladder Mix.

(B) PCR products of purified viral DNA obtained from tagged and untagged hAQP3 constructs by Isopropanol method. Lanes 1 and 6 are PCR product obtained from uninfected insect cell (negative control); 2 and 3 are untagged and tagged hAQP3 viral DNA using Bac primers (1037 bp and 1087 bp); 4 and 5 are untagged and tagged hAQP3 viral DNA using sequence specific primers (882bp); M: 1Kb DNA Ladder Mix.

3.3.4 Expression of AQP4s

The purpose of the study was to verify the expression of the His tagged wheat aquaporins and mammalian AQP3 using Anti His antibody at two time points, 48hr and 72hr infection and also from the frozen infected insect cell pellet at 72hr. To verify the expression of the His tagged wheat and mammalian aquaporins, His tagged La protein was used as a positive control; untagged wheat and mammalian aquaporins was used as a negative control for immunoblot analysis. In an attempt to extract the membrane protein from the aquaporin infected insect cell, integral membrane proteins were phase separated using a non-ionic detergent Triton X-114 and subjected to Ni

116
Chapter 3  

Heterologous expression of AQP genes

affinity pull-down method. All fractions, crude lysate obtained after membrane protein extraction, detergent and aqueous phase of untagged wheat and mammalian aquaporins obtained after phase separation using TritonX-114, IMAC fractions of his tagged aquaporins unbound fraction 1, unbound fraction 2 and Ni\(^+\) bound fraction were ran on the SDS gel to track the aquaporins.

3.3.4.1 Phase separation

The efficiency of the phase separation of the integral membrane proteins were tested using Triton X-114. The aqueous and detergent phase obtained after the phase separation was ran on the SDS gel and Coomassie stained (Figure 3.9).

![Figure 3.9 Phase separation of hydrophilic and amphiphilic proteins.](image)

Cell lysates infected with TaPIP2-9a and hAQP3 viruses were mixed with Triton X-114 and phase separated at 30°C as described under (section 3.2.10).
Aliquots of the diluted and undiluted aqueous and detergent phases were analyzed by 12% SDS PAGE. Lane M – LMW markers.

The comparison of an aqueous and detergent phase of His tagged TaPIP2-9a and His tagged hAQP3 separation shows that the integral membrane proteins have been recovered in the detergent phase (Figure 3.9). This shows that the non-ionic detergent doesn’t bind to the hydrophilic proteins (Makino et al., 1973) and below the temperature of its cloud point Triton X-114 forms micellar solutions and is an efficient solubilizer of membrane proteins (Egan et al., 1976).

3.3.4.2 Expression of wheat and mammalian aquaporin at 48, 72 hr infection and from frozen cell pellet

The expression of wheat and mammalian AQPs were tested at two different time points at 48 and 72 hr infection. Immunoblot analysis didn’t yield any expression for the frozen cell pellet and 48 hr infected His tagged wheat and mammalian AQPs and also for the His tagged La protein, positive control of western blot (data not shown), therefore attempt was made to test the 72 h infected insect cells by phase separation using Triton X-114.

The untagged and tagged wheat and mammalian AQP baculovirus infected the Sf21 cells for 72 h. All the fractions as mentioned above (section 3.3.4) were run on the SDS gel (Figure 3.10, 3.11 and 3.12). 100 mL of the insect cell culture was processed for His tagged La protein, His tagged PIP2-8 and His tagged PIP1-5 samples. Figure 3.10 shows the whole cell lysate, unbound and bound fractions of wheat AQPs along with His tagged La protein to test the efficiency of the IMAC pull-down method and lysate of uninfected insect cells, which was used as a negative control to compare the pattern of the bands obtained with the infected insect cell. Lysates were expected to have many bands as they represent the total protein of the cell and ideally one specific band was expected to be in the Ni⁺ bound fraction. Both His tagged AQP and untagged AQP infected insect cells were phase separated using Triton X-114 at 30°C, whereas the His tagged La protein was not subjected to
Chapter 3  

Heterologus expression of AQP genes

phase separation. The whole cell lysate obtained from the His tagged La protein was subjected to IMAC Ni\(^+\) resin to trap the His tagged La protein. All the protein samples were treated in sample reducing buffer for 1.5 min at 95°C. In figure 3.10 the lanes loaded with whole cell lysate showed many bands as expected, but no substantial reduction in bands were observed in the Ni\(^+\) bound fractions when compared to phase separated sample. The difference in the band pattern between the uninfected and infected insect cells was analysed, but it was quite difficult to differentiate the pattern of bands between the lysate and the Ni\(^+\) bound fractions. The result demonstrates that though the phase separation is efficient in partitioning the aqueous and hydrophilic phases (Figure 3.10) the enrichment of the integral membrane proteins was not achieved when compared to other previous experimental results obtained from the detergent phase by (Bordier, 1981, Pryde and Phillips, 1986, Smythe et al., 1988) using Triton X-114. The IMAC pull down method was also not efficient in concentrating the Histidine tagged proteins. Immunoblot analysis of the same protein samples revealed that the lysate of His tagged La protein and His tagged TaPIP2-9a were expressed at the expected size of \(~40\) kDa and \(~25\) kDa. IMAC –pull-down\(\) fractions of His tagged La and His tagged AQP did not yield any bands on Coomassie blue-stained gels. This may be because of the detergent protein micelle complex and therefore the Histidine tag attached to the protein was not exposed for the antibody to detect or possibly the Histidine tagged protein is not washed out of the Ni\(^+\) - Agarose beads, still sticking on to the Ni\(^+\) beads.
Figure 3.10 SDS PAGE and Western blot analyses of His tagged wheat AQPs, TaPIP2-9a and TaPIP1-2a.
A: SDS PAGE to compare all the fractions of His tagged AQPs. Samples were reduced 1.5 min at 95°C.
B: Western blot analyses of extracts of His tagged Wheat AQPs using anti-Histidine monoclonal antibody. The expression was observed in the lane labelled with bold letters and the band was boxed in black.
3.3.4.3 Concentration of the proteins

The concentration of the protein present in the whole cell lysate fraction of untagged hAQP3 was ~4.5 mg/mL, His tagged hAQP3 is ~7.1 mg/mL, His LA (positive control) was ~9.9 mg/mL and His tagged and untagged wheat AQPs was ~6 mg/ml. Based on Yang et al 1997 who estimated a final yield of AQP4 from the baculovirus expression system of 0.11mg/L of culture it was expected that the 100 ml of culture used in this study would yield a few µg of purified protein only.

3.3.5 Immunoblot analysis

3.3.5.1 Immunoblot analysis of hAQP3

A culture of 200 mL insect cell was infected with the hAQP3 infected recombinant baculoviruses for 72 h. Figure 3.11 A and B shows the coomassie stained gel of the His tagged hAQP3 washed from Ni-Agarose resin, without and with 5mM imidazole, the lanes with bold labeling. The Ni\(^+\) bound fraction washed with 5mM imidazole, did not yield any protein, therefore the fractions washed with 5mM imidazole was not blotted on to the membrane for further analysis. His tagged hAQP3 samples on the figure 3.11 (A) labeled in blue colour was used to verify the expression in the western blot (figure 3.11 C). A strong band was observed at ~25 kDa in the lysate fraction. Many other bands observed in the area of the higher molecular masses may be aggregates or non-specific binding of antibodies.

3.3.5.2 Immunoblot analysis of wheat AQP

Cultures of 200 mL of Sf21 cells were infected with the Histidine tagged and untagged TaPIP2-9a and TaPIP1-2a infected recombinant baculoviruses for 72 h. Figure 3.12 A shows the coomassie stained gel of the His tagged TaPIP2-9a and Figure 3.12 C shows the coomassie stained gel of the His tagged TaPIP1-2a of cell lysates, detergent phase and IMAC resin concentrated samples. His tagged TaPIP2-9a and His tagged TaPIP1-2a labeled in pink and purple colour are the samples applied to another gel for western blot detection of His-tag. A strong band was observed in the Ni\(^+\)
bound fraction of His tagged *TaPIP1-2a* and a faint band in the cell lysate of His tagged *TaPIP1-2a* at ~39 kDa which shows that the His tagged protein was enriched during the IMAC pull down. Many nonspecific bands were observed in the area of the higher molecular masses, which could be glycosylated (Hendriks et al., 2004) forms of the protein despite the sequence of *TaPIP1-2a* not being predicted to have an N-glycosylation site. The glycosylation would cause a slight shift in the mobility on SDS-polyacrylamide gels. The 25 kDa band corresponds with the predicted size of non-glycosylated form of the *TaPIP1-2a* and the presence of a more intense band at 38-40 kDa is possibly indicates oligomerisation or different forms undergoing transport from the ER through the golgi complex (Hendriks et al., 2004). Moreover the band at ~38 kDa is detected as doublet possibly due to glycosylation or proteolysis.

A few faint bands were seen in the Ni$^+$ bound fraction of His tagged *TaPIP2-9a* and no expression was observed in the lysate. Strong expression was not observed for the Ni$^+$ bound fraction of His tagged *TaPIP2-9a* processed from 200 mL culture. Hence with the obtained result, the IMAC pull down was effective for His tagged *TaPIP1-2a* and not for His tagged *TaPIP2-9a* and His tagged hAQP3.
Chapter 3

A

Heterologus expression of AQP genes
Figure 3.11 Concentration of His tagged hAQP3 from 200 mL culture without and with 5% Imidazole verified by a Coomassie-stained gel and immunoblot analysis of His tagged hAQP3.

A and B: Coomassie-stained gel of His tagged hAQP3 and His tagged La protein (positive ctrl), infected for 72 hrs and concentrated via Ni$^+$ Agarose beads without (A) and with using 5% Imidazole (B). C: Immunoblot analysis of His tagged hAQP3 and His tagged La protein (positive ctrl) using monoclonal anti His antibody. The lanes in (A) and (C) labelled in blue colour are the same protein samples, His tagged hAQP3, blotted on to the membrane to verify the expression of the protein. The samples were warmed up for 1.5min at 90°C. Marker – Colour plus prestained.
Chapter 3

Heterologous expression of AQP genes
Chapter 3  

Heterologous expression of AQP genes

**Figure 3.12** Concentration of His tagged *TaPIP2-9a* and His tagged *TaPIP1-2a* from 200 mL culture of 72 h infected insect cells, verified by a Coomassie-stained gel and immunoblot analysis of His tagged hAQP3.

A and C: Coomassie-stained gel of His tagged *TaPIP2-9a* and His tagged *TaPIP1-2a* along with His tagged La protein (positive ctrl), infected for 72 hrs was phase separated using Triton X-114 and concentrated via IMAC resin. The lanes labelled in pink and purple colour are the His tagged *TaPIP2-9a* (A) and His tagged *TaPIP1-2a* (C) are protein samples run in another gel blotted on to the membrane shown in B.

B: Immunoblot analysis of His tagged *TaPIP2-9a* and His tagged *TaPIP1-2a* along with His tagged -La protein (positive ctrl), using monoclonal anti His antibody. The lanes labelled in pink colour are His tagged *TaPIP2-9a* and purple colour is His tagged *TaPIP1-2a*. Marker – Colour plus prestained marker.

**Note:** The crude extract samples of his hAQP3, untagged hAQP3, HisLa and Uninfected insect culture were diluted in 1:10 dilution using Sample Reducing buffer. The fractions of His hAQP3 were diluted in 1:2 using SRB. The samples were heated at 95°C for one and half min. 10 µl of all crude samples including uninfected cell culture and fractions 1 and 2 were loaded for both, staining and Western Blot gels. For staining gel, 5 µl of affinity purified samples of HisLa and his hAQP3 were loaded while for Western Blot, 15µl of both these samples were loaded.
3.4 Summary

In summary the subcloning of the Histidine tagged and untagged versions of wheat and mammalian AQPs was successful and verified by restriction digestion using restriction enzymes and the orientation was confirmed by sequencing using vector based bac primers. The recombinant baculovirus was verified by PCR on viral DNA using sequence specific primers and bac primers.

Membrane protein was extracted (from whole cell lysates) from both His tagged and untagged wheat and mammalian AQPs. The aqueous phase and the detergent phase was separated using non-ionic detergent Triton X-114. The detergent phase with the His tagged membrane proteins were concentrated using IMAC Ni\(^{+}\) resin. The greatest expression of recombinant AQP proteins was achieved at 72 h infection. The efficiency of the phase separation was tested. This study demonstrates that the phase separation is efficient by separating the two phases in the 12%SDS gel (Fig 4.11). The SDS results obtain from the His tagged LA (positive control), untagged and His tagged versions of wheat and mammalian AQPs were not able to demonstrate expression. The whole cell lysate had many bands which represents all the proteins, phase separated sample, all fractions (unbound and bound) from the Ni\(^{+}\) agarose resin has many proteins instead one single protein at ~25 to ~30 kDa. The use of 5% imidazole during the wash of His tagged hAQP3 did not yield any proteins. Hence the results were inconclusive with the SDS results. Immunoblot analysis revealed that the positive control His-La protein was expressed at ~42 kDa and His tagged hAQP3 at ~25 kDa in the whole cell lysate with many bands in the higher molecular mass. These higher molecular weight bands might represent the glycosylated form of the protein or could be the degradation of the protein due to the lack of protease inhibitor or could occur due to cross reactivity. Immunoblot analysis of His tagged TaPIP1-2a (Fig 4.13 B) showed up expression at ~25 kDa in the Ni\(^{+}\) bound fraction and at ~40 kDa and above, in both the cell lysate and in Ni\(^{+}\) bound fraction, this could be a potential dimer formation. Faint bands were observed for His tagged TaPIP2-9a (Fig 4.13 B) in the Ni\(^{+}\) bound fraction, but a strong
Chapter 3  

Heterologous expression of AQP genes expression was obtained in the cell lysate of His tagged *TaPIP2-9a* (Fig 4.10 B). Therefore the expression of the His tagged AQPs and the His tagged LA protein was shown, but optimization of the experimental protocols is required for total protein extraction, phase separation and affinity purification using Ni+ affinity chromatography.
4

Functional testing of aquaporins expressed in insect cells
4 Functional testing of aquaporins expressed in insect cells

Extraction, purification and reconstitution of putative aquaporins into membrane vesicles to enable stop flow spectrophotometry is difficult, expensive and time consuming. This chapter describes an attempt to develop a novel method to assay the activity of aquaporin proteins by measuring the change in size of the whole insect cell after subjecting the cell to hypo-osmotic shock.

4.1 Introduction

4.1.1 Glycerol transport

Some aquaporin family members, referred to as aquaglyceroporins, transport glycerol and other uncharged solutes in addition to water. The aquaporins can be distinguished from aquaglyceroporins based on primary amino acid sequence alignments (Borgnia et al., 1999). Mammalian AQP3 was the first aquaglyceropporin to be cloned and exhibits water, glycerol and urea permeability (Echevarria et al., 1994, Ishibashi et al., 1994). Other mammalian aquaporins AQP7 and AQP10 also transport water, glycerol and urea when expressed in Xenopus oocytes (Ishibashi et al., 1997, Ishibashi et al., 2002). AQP9 transports water, glycerol and urea, but also is permeable to a wide range of other solutes in oocytes (Tsukaguchi et al., 1998).

4.1.2 Role of glycerol in mammals and plants

Glycerol is used as a carbon source for glycolysis and for lipid biogenesis. Two types of mechanisms are involved in transporting glycerol into the cytoplasm: passive diffusion across the lipid bilayer, or facilitated uptake when the external glycerol concentration is low (Richey and Lin, 1972). Hepatocytes in mammals are able to take up blood glycerol for gluconeogenesis; however this is unlikely to be an important pathway in plants. Glycerol is more likely to be transported across tonoplast membranes
as it is an osmoprotectant (Aldesuquy et al., 2012) and the aquaporin class known as NIPs have been identified as glycerol transporters.

### 4.1.3 Plant Aquaglyceroporins

In plants so far *AtNIP6-1* (GenBank: AEE36445.1) was identified to transport glycerol and function as a membrane channel that selectively transports water, small neutral molecules, and ions out of and between cells.

Tobacco NtAQP1, a PIP1 member transports glycerol (Biela et al., 1999, Gerbeau et al., 1999), XIPs a new member of MIPs identified in the genomes of plants and fungi also transports glycerol (Bienert et al., 2011). Other Plant NIPs in tobacco demonstrated transport of glycerol in addition to water *AtNIP1-1* and *AtNIP1-2* (Weig and Jakob, 2000). Aquaporins are tetrameric in the membrane whereas the glycerol transporters exist as monomers (Lagrée et al., 1999). The physiological evidence for a MIP-mediated glycerol transport has not been shown in plants, though there were several isoforms of MIP subfamilies conducted glycerol (Bienert et al., 2011). The putative primary amino acid sequences of the *TaPIP2-9a* and *TaPIP1-2a* in comparison with other glycerol transporting bacterial and mammalian aquaporins show that the two full length wheat aquaporins obtained from this study shows that these aquaporins could potentially be a non glycerol transporter (refer chapter 2, figure 2.15).

### 4.1.4 Aquaglyceroporins

The known mammalian aquaglyceroporins are hAQP3, hAQP7, hAQP9 and hAQP10. In this study hAQP3 was used as positive control to test the function of the unknown wheat AQPs. hAQP3 is an aquaglyceroporin, transports water at physiological pH and glycerol channel at pH values around 6.1, when expressed in *Xenopus laevis* oocytes (Zeuthen and Klaerke, 1999).
Chapter 4  
Functional testing of Aquaporins

<table>
<thead>
<tr>
<th>Mammalian Aquaglyceroporins</th>
<th>GenBank accession numbers</th>
<th>Experimental evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>hAQp3</td>
<td>AB001325.1</td>
<td>Yes (Ishibashi et al., 1995)</td>
</tr>
<tr>
<td>hAQp7</td>
<td>AB006190.1</td>
<td>Yes (Kuriyama et al., 1997)</td>
</tr>
<tr>
<td>hAQp9</td>
<td>AB008775.1</td>
<td>Yes (Ishibashi et al., 1998)</td>
</tr>
<tr>
<td>hAQp10</td>
<td>AB066105</td>
<td>Yes (Ishibashi et al., 2002)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E.coli</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GifP</td>
<td>AAA23886.1</td>
<td>Yes (Weissenborn et al., 1992)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plant Aquaglyceroporins</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AtNIP6-1</td>
<td>AEE36445.1</td>
<td>Yes (Theologis et al., 2000)</td>
</tr>
</tbody>
</table>

| Rice aquaporin             | Unknown                  | No                    |
| wheat Aquaporin            | Unknown                  | No                    |
| Barley Aquaporin           | Unknown                  | No                    |

Table 4.1 Aquaglyceroporins in mammals and plants.

GenBank accession numbers and functions of the existing mammalian and plant aquaglyceroporins.

4.1.5 Existing methods to test the function of the aquaporins: advantages and disadvantages

The functions of the AQPs are currently tested by video or time lapse image analysis when expressed in Xenopus oocytes (Zhang and Verkman, 1991, Zhang et al., 2008, Chaumont et al., 2001, Katsuhara et al., 2002, Kammerloher et al., 1994, Biela et al., 1999, Ishibashi et al., 1997, Ishibashi et al., 2002) and light scattering method when expressed in yeast (Ishikawa et al., 2005, Laizé et al., 1995, Coury et al., 1998, Daniels et al., 2006, Suga and Maeshima, 2004) or when purified protein is reconstituted into artificial vesicle. The existing light scattering methods require pure protein to prepare vesicles, electron microscopy to check the uniformity and mean size of the vesicles, stop flow spectrometry to measure the rate of water transport per second, which is expensive and takes time.
4.1.6 Proposed method in this study: advantages and disadvantages

The aim of this work was to test the water transport activity of wheat aquaporins (tagged and untagged) using the whole insect cell. If the method proved to be quantitative it could be used as a screening method to test the function of the aquaporin using whole insect cells. The method would be cheaper compared to existing methods and could be performed in laboratories that do not have access to the specialised equipment required for existing methods.

4.1.7 Cell volume as a measure of water transport

Plasma membranes are more permeable to water than to most other small molecules, ions and macromolecules. This permeability is due partly to simple diffusion of water through the lipid bilayer and partly to protein channels (aquaporins) in the membrane that selectively permit the passage of water. In hypertonic solution (higher osmolarity outside the cell than the cytosol) the cell shrinks as water flows out. In hypotonic solution (lower osmolarity outside the cell than the cytosol) the cell swells as water flows in and if, unsupported by cell wall, eventually bursts. Cells generally contain higher concentration of biomolecules and ions than their surroundings, so osmotic pressure tends to drive water into the cells. The inward movement of water would distend the plasma membrane and eventually cause bursting of the cell (osmotic lysis).

There are three mechanisms to prevent cell lysis. The plasma membrane of plants and bacteria is made up of nonexpandable cell wall to prevent osmotic lysis and also to resist osmotic pressure. The presence of contractile vacuole, pumps the water out of the cell in certain fresh waer protozoans which live in a hypotonic medium, whereas in multicellular animals the osmolarity of blood plasma and interstitial fluid are maintained close to that of cytosol. In plants the mechanical rigidity is achieved by the osmotic pressure.
Chapter 4 Functional testing of Aquaporins

The existing methods of measuring water transport rates use the change in size of intact oocytes, cells or membrane vesicles in response to a change in the external osmotic potential.

In oocytes the half life of water transport is approximately 30 seconds whereas in vesicles it is in the order of milliseconds in keeping with a size for the oocyte of 1 mm and less than one micron for vesicles. Insect cells at 50 microns in diameter are expected to give an intermediate time course for osmotic swelling.

4.2 Materials and Methods

4.2.1 Experimental procedures

Swelling Assays— swelling assays were performed by the transfer of the insect cell from isoosmotic to hypo-osmotic (100 mosmol/kg) solution. After the transfer of the cell, volume changes were recorded with Olympus camera (Olympus, Japan) on Olympus Eclipse E 300 inverted microscope (CKX41, Olympus, Japan). Images were captured at intervals of 30 s for 10 min and digitized by the ImageJ software (Version 1.44, http://rsbweb.nih.gov/ij/download.html). The data were analyzed as proportional change in volume and normalized to the initial volume at time 0. All the assays were performed at room temperature (25 °C).

4.2.2 Method development

The osmotic water permeability of the Sf21 cells infected with human AQP3 (hAQP3), wheat aquaporin TaPIP2-9a and TaPIP1-2a, Beta-galactosidase, GFP and uninfected insect cells were measured using whole insect cells. After several trial and errors, the proposed method was used successfully employed for the glycerol experiment.

4.2.2.1 Water, Glycerol and Mercury experiment triplicates

The T-flask containing insect cells infected with untagged constructs for 72h were tested for the water, glycerol and mercury transport activity. The freshly seeded uninfected insect cells at the density of 2.5*10^5 cells/mL in the
25cm$^2$ T-flask were allowed to settle for 20 min in the room temperature and tested for the water, glycerol and mercury transport activity by subjecting the cells to hypo-osmotic shock and images were captured for every 1min until 10min (graph was plotted from 0 min till 2 mins).

The infected insect cells were seeded at the density of 2.5*10$^5$ cells/mL and maintained in the 25cm$^2$ T-flask for 72hours in the total volume of 5ml. The cells were subjected to hypo-osmotic shock by replacing 4 ml of insect cell by 2 mL of sterile water using 10 mL serological pipette (Castor ®STRIPETTE) for water experiment, 160 mM glycerol for glycerol experiment and 0.3 mM mercury for mercury experiment. Each experiment was performed independently in triplicate. The cells subjected to hypo-osmotic shock were viewed under microscope, area of focus was randomly chosen and an image was captured every 30sec until 10min. The images were processed using ImageJ software, cross-sectional area was measured and cell volume calculated and graphed.

4.2.2.2 Image processing

4.2.2.2.1 Measuring the insect cells size using ImageJ

ImageJ software was used to measure the change in the size of the insect cells subjected to hypo-osmotic pressure. Individual cells in each image were manually selected in ImageJ (rejecting the few dead and misshapen cells) and the area of each cell in the image was measured (refer appendix 8 for the protocol).

The area in pixel square (px$^2$) was converted to area in micron square (µm$^2$). Cross-sectional area of the cell was converted to volume of the cell (Cm$^3$) and then absolute volume change (µm$^3$) was calculated for all the time points.
4.2.3 MATLAB

Matlab® 2011a is computational software with built-in standard functions allowing the user to process data and produces a mathematical solution. The functionality of Matlab used for this research is to approximate an exponential trend line to a given set of data points.

To calculate the osmotic water permeability coefficient ($P_f$) values the absolute volume change data was fitted into a single and a double exponential function expressed in equations (1) and (2). Derivation of these functions from first principles are well described in (van Heeswijk and van Os, 1986):

$$ S = S_\infty - \Delta S \exp(-K_e t) \quad (1) $$

and

$$ S = S_\infty - \Delta S_s \exp(-K_s t) - \Delta S_f \exp(K_f t), \quad (2) $$

where $\Delta S_s$ and $\Delta S_f$ correspond to the signal changes with either a slow rate constant ($K_s$) or fast rate constant ($K_f$). $K_e$ is an approximate rate constant (van Heeswijk and van Os, 1986).

The average rate constant $K_{de}$ can be determined from $\Delta S = \Delta S_s + \Delta S_f$, 

$$ K_{de} = \frac{K_s \Delta S_s + K_f \Delta S_f}{\Delta S} $$
4.3 Results and Discussion

4.3.1 Hypo-osmotic water transport and mercury sensitivity

The biological triplicate data obtained from the newly developed method in terms of pixel square (px²) was converted to absolute volume change and fitted in the double/single exponential curve using Matlab (data not shown). The changes in cell volume in response to hypo-osmotic shock for uninfected cells and cells infected with recombinant baculoviruses expressing aquaporin TaPIP2-9 are shown in figure 4.1 A and B (upper curves).

The volume of the uninfected insect cells and infected insect cells increased after cells were subjected to hypo-osmotic medium of half normal osmotic strength. The final equilibrium volume showed no clear trends with type of virus and uninfected cells showed the greatest increment (data not shown). However the absence of data in the first 30 seconds did not enable the calculation of a zero time rate of change in volume. However, based on the first time point (30 sec) the presence of mercury reduced the rate of change in volume for uninfected control cells and cells infected with PIP2-9 virus (Fig 4.1 A, B and C lower curves).

The effect of mercury on the volume change at this time point for all recombinant viruses is shown in Fig 4.1 D. The overall trend exhibited by the uninfected insect cells and insect cells infected with viruses shows increase in water transport activity at the initial time point 30 sec (0.5 min) in response to hypo-osmotic shock except for beta-galactosidase and TaPIP1-2a.

The uninfected and the infected insect cells were also tested for mercury sensitivity in the presence of water. The uninfected insect cells i.e the endogenous aquaporins present in the insect cells were highly sensitive to mercury. The negative control GFP, aquaporins hAQP3 (positive control) and TaPIP2-9a exhibits significant change in mercury treatment. This method also demonstrates that the known mercury sensitive mammalian AQP3 (Kuwahara et
al., 1997) is sensitive to mercury. Hence this method could also be used to differentiate the mercury sensitive and insensitive aquaporins. TaPIP1-2a and beta-galactosidase infected insect cells exhibit increase in water transport activity in the presence of mercury and did not show mercury sensitivity. Although the primary amino acid sequence of TaPIP1-2a and TaPIP2-9a shows predicted mercury sensitive sites (refer appendix 3), the proposed method demonstrates that the TaPIP1-2a is mercury insensitive. Therefore further test using stop flow spectrophotometry could be conducted to determine the rate of the water transport.

Attempts to acquire data that would yield zero time rate estimates were not successful due to technological problems in capturing and focussing the cells within 0.5 minute, therefore the rate of water transport was unable to be obtained. The kinetics of the water swelling was too rapid to be able to acquire data that yields zero-time rate. Attempts to fit double-exponential functions to the date were not successful for water transport even in the presence of mercury.
Chapter 4 Functional testing of Aquaporins

A

B
Figure 4.1 Volume changes of insect cells in response to hypo-osmotic shock.
Volume changes of insect cell infected with recombinant baculoviruses in response to hypo-osmotic change (in the presence of water and Mercury), A: insect cells infected with GFP; B: insect cells infected with TaPIP2-9a; C: insect cells infected with TaPIP2-9 (all data is from triplicates) and D: Bar graph representing the change in volume of the insect cells at 0.5 min in response to hypo-osmotic shock. Data in A, B and C are expressed as the mean ± SE, with the number of replicates shown in the figure.
UIC - uninfected insect cell, GFP- Green fluorescent protein, Betagal – Beta-galactosidase. Values are µm³.
4.3.2 Isoosmotic glycerol transport

The volume of the uninfected insect cells and infected insect cells increased after cells were subjected to iso-osmotic medium containing glycerol at 160 mM. The time taken for the uninfected and infected insect cells to reach equilibrium was greater than 10 mins. The effect of glycerol on the volume change at this time point for all recombinant viruses is shown in Fig 4.2 except for TaPIP2-9a. The overall trend exhibited by the uninfected insect cells and infected insect cells shows gradual increase in the volume of the cell. Attempt to fit the triplicate data obtained from glycerol in the double/single exponential function was unsuccessful (data not shown), so the rate of glycerol transport was not calculated.

![Glycerol EXPT](image)

**Figure 4.2 Volume changes of insect cells in response to isoosmotic shock.** Volume changes of insect cell infected with recombinant baculoviruses in response to isotonic medium with 160mM glycerol.
4.3.3 Statistical analysis

All of the treatments were applied in triplicate (where mentioned duplicates, in case of contamination). The data were expressed as mean ± standard error of the mean (SEM). The results of the water, mercury and glycerol treated insect cells were analyzed with a one-way ANOVA test, compared to cell control (UIC) and virus controls (beta-galactosidase and GFP).

*Water treatment*

The virus controls beta-galactosidase and GFP showed a P value < 0.05. None of the aquaporin infected insect cells showed a significant P value.

*Mercury treatment*

Statistical analysis shows that data obtained for GFP, *TaPIP2-9a* and *TaPIP1-2a* in the mercury treatment is significant with P value < 0.05 and not significant for the beta-galactosidase, hAQP3 and UIC.

*Glycerol treatment*

In the glycerol treatment beta-galactosidase and hAQP3 yeilded a significant statistical P value < 0.05.

Based on the statistical values it could be stated that the test aquaporins *TaPIP1-2a* and *TaPIP2-9a*; aquaporin control hAQP3 transports water faster and reaches equilibrium in a minute, hence did not yield a significant statistical value. *TaPIP2-9a* could potentially be a mercury sensitive protein; though hAQP3 shows mercury sensitivity (Figure 4.1 D) it is not statistically significant. The statistically significant p value of *TaPIP1-2a* may likely be a mercury insensitive protein; further experiments have to be performed to satisfy this statement. The known aquaglyceroporins hAQP3 further proves that this is a glycerol transporter. Though the test aquaporins *TaPIP1-2a* and *TaPIP2-9a*; virus controls beta-galactosidase and GFP and the cell control UIC exhibits similar trend (Figure 4.2), the statistical analysis showed that the already known aquaglyceroporins hAQP3 yeilded a P value < 0.05. The new method has the
potential to differentiate the glycerol transporter and non-glycerol transporter; the results obtained from this study could be reconfirmed by repeating the experiment in the stopflow spectrophotometer to further validate the low cost new method.

4.4 Summary

In conclusion our results constitute the first report on the functional testing of human and wheat AQPs using whole insect cell. Although the water transport occurred in the infected and uninfected insect cells, due to technological problems and the fast reaction kinetics the rate of water transport was unable to be calculated. However the rate of water transport of the whole insect cells infected with TaPIP1-2a under glycerol exposure and mercury treatment was monitored and measured.

The data obtained for the water, glycerol and mercury treatment for the insect cells infected with human AQP3, TaPIP2-9a, beta-galactosidase, GFP and uninfected insect cells did not fit in both the single and double exponential graph. Though the water transport reaction occurred in the insect cells, it was too fast for the proposed method to monitor the changes in the cell volume. Measurement of water permeability in the intact Sf21 cells by cell swelling assay was unsuccessful because of the time constraints.

Analysis of the primary amino acid sequence of TaPIP1 and TaPIP2 with other known water transport, glycerol transport and mercury sensitive aquaporins demonstrates that the two wheat aquaporins TaPIP1-2a and TaPIP2-9a may potentially transports water and could be mercury sensitive aquaporins, bearing mercury sensitive amino acid residues at the respective positions (appendix 3).
5

General conclusions and future direction
5 General conclusions and future direction

5.1 Conclusions and contribution from this work

The principal aim of this thesis was to express wheat aquaporins in the baculovirus insect cell system, extract membrane protein using TritonX-114 and verify the expression using anti-His antibody. A novel test was devised for the function of the aquaporin protein using a whole insect cell swelling assay. In order to achieve the above aims, this thesis reviews the existing data of wheat aquaporins (Chapter 1, table 2), RT-PCR of select wheat PIP genes (chapter 2, section 2.3.3.1.3 and 2.3.3.2.3), subcloning of the select wheat PIP genes in the untagged and tagged baculoviral transfer vectors, phase separation of membrane protein using the detergent Triton X-114, verification of the expression of the Histidine tagged aquaporins and functional testing of aquaporins using the newly developed method. The results from this work demonstrated that the aquaporins are highly expressed in the early seedling stage compared to the other stages. The phase separation is effective in separating the aqueous and detergent phase but the affinity purification was not effective in separating the His tagged proteins and thus the immunoblotting was interfered with many nonspecific bands. The new method employed to test the function of the wheat and mammalian aquaporins was successful for the slow water transporting proteins (for eg, The PIP1 aquaporins in plants transports less/slow water). Therefore the attempt to phase separate the membrane protein and whole insect cell swelling assay can be used as a preliminary method to study the function of the protein which would greatly reduce the cost and time. On the hindside there are technical limitations for the fast water transporting protein, hence to obtain qualitative data stop flow experiment has to be performed.

Chapter 2: The present study amplified two full length wheat AQP genes TaPIP1-2a cds (867 bp) and TaPIP2-9a cds (897 bp) from the early seedling stage (8-day-old) root cDNA of the common wheat cultivar Cranbrook. Further one isoform of TaPIP1-2 (EU177547) gene represented as TaPIP1-2a, one isoforms of TaPIP1-6 (EU177551) gene represented as TaPIP1-6a and
two isoforms of \( TaPIP2-9 \) (EU177562) gene represented as \( TaPIP2-9a, \) \( TaPIP2-9b \) isolated from genomic DNA of \( T.aestivum \) cv.Cranbrook and the gene structure was constructed by intron/exon nucleotide sequence comparison. Like all other AQPs the \( TaPIP1-2a \) cds and \( TaPIP2-9a \) cds have six conserved transmembrane domains (chapter 2, fig 2.5) and further the putative protein sequence comparision of the two full length wheat cds along with rice AQPs demonstrates that the residues differentiates PIP1 and PIP2 in plant AQPs are conserved (refer appendix 3). The location of the genes \( TaPIP2-9a \) (gDNA) and \( TaPIP2-9b \) (gDNA), \( TaPIP1-2a \) (gDNA) and \( TaPIP1-6a \) (gDNA) in the hexaploid wheat genome could potential be AA/BB because these genes exhibit 98% and 92% identity with the existing full length wheat AQPs in the database (Refer table 2.11). The results from this work will be published.

In chapter 3 the attempt of subcloning these two full length wheat AQP genes, positive control hAQP3 in the Baculoviral vectors pBacPAK9 (untagged) and pBacPAKHis1 (tagged) were successful and the orientation was verified by sequencing. In total six constructs (both untagged and tagged) were constructed. The insect cells infected with Histidine tagged AQP constructs were used for membrane protein extraction and western blotting. The insect cells infected with untagged AQP constructs were used to test the function of the protein by subjecting the cells to the hypo-osmotic medium. The Beta-galactosidase and Green fluorescent protein (GFP) were the virus controls (negative control) for the experiment. Immunoblot analysis to verify the expression of the recombinant aquaporins at 24 h, 48 h and 72 h infection revealed that the greatest expression was observed at 72 h infection, where Histidine tagged LA protein was used as a positive control. Analysis of the primary amino acid sequence for the substrate selectivity and mercury sensitivity with the known human and plant aquaporins demonstrates that the two wheat aquaporins \( TaPIP2-9a \) cds, \( TaPIP1-2a \) cds may transport only water and not glycerol and could potentially be a non ammonia transporter.
In chapter 4 the results obtained from the water, glycerol and mercury experiment using the whole insect cell swelling assay shows that the osmotic driven water transport and glycerol transport occurs. Under hypo-osmotic condition the water transport is very fast and reaches equilibrium in a minute, whereas in the isosmotic condition the insect cell reaches the final equilibrium in 10 mins. But due to technical limitations the new method did not yield rate of water transport for the wheat aquaporins and mammalian aquaporin. Hence the water transport activity of the two wheat aquaporins was unable to be compared with the literature to determine the significance of these two proteins in the physiological role of wheat. Nonetheless the new method could be used as the screening test for the water transport activity and mercury sensitivity. These results were presented at the “Plant Membrane Biology, 15th International Workshop” Adelaide 2010.

5.2 Future direction

The result of chapter 3 demonstrates that the Histidine tagged AQPs expression can be verified by the phase separation method and that wheat aquaporins can be expressed in the baculovirus system. Although the wheat and mammalian AQPs are expressed in the cell lysate or in the Ni\(^{2+}\) bound fraction in the different attempts, these studies remain incomplete. The following research directions are suggested to optimize the membrane protein extraction and phase separation of wheat aquaporins.

- Existing wheat aquaporin genes can be compared with rice micro array database and functionally important genes can be identified and those genes can be amplified using RACE PCR and the functions can be tested using heterologous expression system.

- The chromosomal location of wheat aquaporins obtained from this study TaPIP2-9a cds, TaPIP1-2a cds and likely new homeologs of TaPIP1-2, TaPIP1-6 and TaPIP2-9 (TaPIP2-9a and TaPIP2-9b) can be located using nullisomic tetrasomic lines.
Chapter 5 General result and discussion

- The tissue localisation of the wheat *TaPIP1-2* and *TaPIP2-9* genes can be investigated by designing sequence specific oligos, using the DNA sequence information obtained from this study.

- Existing wheat aquaporin genes can be compared with rice micro array database and functionally important genes can be identified and those genes can be amplified using RACE PCR and the functions can be tested using heterologous expression system.

- Sucrose gradient ultracentrifugation can be used instead of Sorvall centrifugation for the effective fractionation (Yakata *et al.*, 2011).

- Excess detergent bound to the integral membrane protein must be removed by equilibrium dialysis, affinity chromatography; hydrophobic interaction chromatography, ion-exchange chromatography, precipitation using acetone and ultrafiltration (Simith, 2011) and detergent/ lipid ratio should be calculated.

- High concentration of imidazole (10mM) can be used to reduce non-specific binding of the His tag from the Ni+ affinity chromatography.

- The predicted protein sequence of the full length *TaPIP1-2a* cds and *TaPIP2-9a* cds could also be used to design protein specific peptide to raise antibodies for the immnoblots.

- The results obtained from the functional testing of aquaporins (chapter 4) can also be tested by carrying out oocyte swelling assays, commonly used method to test the aquaporin function.
6
References
6 References


References
References


References


aquaporin and H+-ATPase related to motor cell function in Mimosa pudica L. *Plant Physiology*, 114, 827-834.


References


References


References


PRYDE, J. G. & PHILLIPS, J. H. 1986. FRACTIONATION OF MEMBRANE-PROTEINS BY TEMPERATURE-INDUCED PHASE-SEPARATION IN TRITON X-114 -


References


167
References


7

Appendices
# Appendices

## Appendix 1 Alignment of *TaPIP1-2a* full length coding sequence (cds), *TaPIP1-2a* and *TaPIP1-2* (gDNA) to predict intron and exon junction.

<table>
<thead>
<tr>
<th>TaPIP1-2a cds</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-2a gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>TaPIP1-2 gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TaPIP1-2a cds</th>
<th>110</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
<th>160</th>
<th>170</th>
<th>180</th>
<th>190</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-2a gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>TaPIP1-2 gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TaPIP1-2a cds</th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
<th>250</th>
<th>260</th>
<th>270</th>
<th>280</th>
<th>290</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-2a gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>TaPIP1-2 gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TaPIP1-2a cds</th>
<th>310</th>
<th>320</th>
<th>330</th>
<th>340</th>
<th>350</th>
<th>360</th>
<th>370</th>
<th>380</th>
<th>390</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-2a gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>TaPIP1-2 gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TaPIP1-2a cds</th>
<th>410</th>
<th>420</th>
<th>430</th>
<th>440</th>
<th>450</th>
<th>460</th>
<th>470</th>
<th>480</th>
<th>490</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-2a gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>TaPIP1-2 gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TaPIP1-2a cds</th>
<th>510</th>
<th>520</th>
<th>530</th>
<th>540</th>
<th>550</th>
<th>560</th>
<th>570</th>
<th>580</th>
<th>590</th>
<th>600</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP1-2a gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>TaPIP1-2 gDNA</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

---

*Oryza s*

---

*Triticum*

---

*Hordeum*

---

*Triticum*

---

*Hordeum*
The intron and exon junction of TaPIP1-2a coding sequence, TaPIP1-2a and TaPIP1-2 (EU177547) genomic sequences were predicted by comparing the nucleotide sequence. The regions highlighted in black are intron regions, unshaded regions are exons and the region highlighted in ash denotes the partial gene in the database.

Alignment created in ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and displayed in Bioedit (http://www.mbio.ncsu.edu/bioedit/bioedit.html).
Appendix 2 Alignment of TaPIP2-9 gene obtained from cDNA and genomic DNA to assign intron and exon.
Appendices
<table>
<thead>
<tr>
<th>TaPIP2 Ra1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TaPIP2-9a cds</th>
<th>GGTACCCCACGGGCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaPIP2-9a AS</td>
<td>----------------</td>
</tr>
<tr>
<td>TaPIP2-9a gDNA</td>
<td>ACGGG.T---------</td>
</tr>
<tr>
<td>TaPIP2-9b gDNA</td>
<td>----------------</td>
</tr>
<tr>
<td>TaPIP2-9 gDNA</td>
<td>---------------</td>
</tr>
</tbody>
</table>

Alignment of the homeologous *TaPIP2-9* and *TaPIP2-9b* genes obtained from the genomic DNA of wheat, known partial *TaPIP2-9* gene from the database (EU177562), coding sequence of *TaPIP2-9a* cds and alternate splice variant of *TaPIP2-9a AS* were aligned to predict the intron exon junction and also to compare the similarities between the exon regions. A shading bases in the known partial *TaPIP2-9* gene represents the unidentified N and C-terminal regions represented in “-”, the regions highlighted in black in the splice variant denotes the lack of expressed region. The stop codon (TGA) is underlined. Arrow denotes the primer positions. **Alignment created in ClustalW** (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and **displayed in Bioedit** (http://www.mbio.ncsu.edu/bioedit/bioedit.html).
### Appendix 3: ClustalW Alignment of Deduced Amino Acid Sequence of *TaPIP1-2a* cds and *TaPIP2-9a* cds with Rice MIPs

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TaPIP1-2a</strong></td>
<td>MEGKE-EDVR LGANKYSERQ PIGTAAQGS- EDKDYKE--- PPPAPLFEPG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP1;1</strong></td>
<td>MEGKE-EDVR LGANRYSERQ PIGTAAQGAG DDKDYKE--- PPPAPLFEPG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP1;2</strong></td>
<td>MSGWEPTSSR RGSRSARRRR APTTRTTRSR RRRRCSSQGS SSRGPSAQG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP1;3</strong></td>
<td>MEGKE-EDVR LGANRYTERQ PIGTAAQGA- EEKDYRE--- PPAAPVFEVE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TaPIP2-9a</strong></td>
<td>-------------- --MA--KEVS EEPEH---AAP AHKDYSD--- PPPAPLFDGM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;1</strong></td>
<td>-------------- --MG--KDEV MESGAAGEF AAKDYTD--- PPPAPLIDAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;2</strong></td>
<td>-------------- --MA--KDIE ASAP--EGGEF SAKDYTD--- PPPAPLIDVE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;3</strong></td>
<td>-------------- --MA--KDIE AAAAAEGGEY MAKDYSD--- PPPAPLIDAE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;4</strong></td>
<td>-------------- --MG--KEVD VSTLEAGG-- ARDYID--- PPPAPLVVDV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;5</strong></td>
<td>-------------- --MG--KEVD VSTLEAGG-- VRDYED--- PPPAPLVDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;6</strong></td>
<td>-------------- --MS--KEVS EEPEH--VRP --KD YT--- PPPAPFLDVG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;8</strong></td>
<td>-------------- --------------- -MAAGSGSGS NPDKYQD--- PPPAPLVDTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;7</strong></td>
<td>-------------- --MASKEEVA VETVEGAAA AKAPYWD--- PPPAPLDT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TaPIP1-2a</strong></td>
<td>ELKSWSFYRA GIAEFMATLL FLYVTILTV GYSGAASK-- ------CATV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP1;1</strong></td>
<td>ELKSWSFYRA GIAEVTPLL FLYITLTV GYKSSESSK-- ------CATV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP1;2</strong></td>
<td>SPSWPSSSS -------------- -SYITALNAM GVNNSTSK--- CATV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP1;3</strong></td>
<td>ELTSWSFYRA GIAEFVATFL FLYISILTV GVNSASK--- ------CATV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TaPIP2-9a</strong></td>
<td>ELRMWSFYRA GIAEFVATLL FLYITVATV YGKYQ-SAAD P-----CGGV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;1</strong></td>
<td>ELGWSLYRA VIAEHIATLL FLYVATV VYKQTDAAS SGADAACG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;2</strong></td>
<td>ELTKWSLYRA VIAEHIATLL FLYVATV VYKHSADTV NTDDACSG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;3</strong></td>
<td>ELTKWSLYRA VIAEHIATLL FLYVATV VYKHQSDPQA NAADACSG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;4</strong></td>
<td>ELGKWSLYRA LIAEFVATLL FLYVTATV VYKQTDAAV NGADAACG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;5</strong></td>
<td>ELKWSLYRA VIAEHIATLL FLYVATV -------------- -CGGV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;6</strong></td>
<td>ELKWSLYRA LIAEFVATLL FLYVATV VYKQV-SAAD Q-----CGHV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;8</strong></td>
<td>ELGKWSLYRA AIAEFVATLL LVCISVSTVI GKYQRSGE--- ------CGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OsPIP2;7</strong></td>
<td>ELKWSLYRA LIAEFMATLL FLYVSATV VYKNQRADV A-----CTGV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TaPIP1-2a  |
OsPIP1;1  |
OsPIP1;2  |
OsPIP1;3  |
TaPIP2-9a  |
OsPIP2;1  |
OsPIP2;2  |
OsPIP2;3  |
OsPIP2;4  |
OsPIP2;5  |
OsPIP2;6  |
OsPIP2;8  |
OsPIP2;7  |

TaPIP1-2a  |
OsPIP1;1  |
OsPIP1;2  |
OsPIP1;3  |
TaPIP2-9a  |
OsPIP2;1  |
OsPIP2;2  |
OsPIP2;3  |
OsPIP2;4  |
OsPIP2;5  |
OsPIP2;6  |
OsPIP2;8  |
OsPIP2;7  |

TaPIP1-2a  |
OsPIP1;1  |
OsPIP1;2  |
OsPIP1;3  |
TaPIP2-9a  |
OsPIP2;1  |
OsPIP2;2  |
OsPIP2;3  |
OsPIP2;4  |
OsPIP2;5  |
OsPIP2;6  |
OsPIP2;8  |
OsPIP2;7  |

TaPIP1-2a  |
OsPIP1;1  |
OsPIP1;2  |
OsPIP1;3  |
TaPIP2-9a  |
OsPIP2;1  |
OsPIP2;2  |
OsPIP2;3  |
OsPIP2;4  |
OsPIP2;5  |
OsPIP2;6  |
OsPIP2;8  |
OsPIP2;7  |

TaPIP1-2a  |
OsPIP1;1  |
OsPIP1;2  |
OsPIP1;3  |
TaPIP2-9a  |
OsPIP2;1  |
OsPIP2;2  |
OsPIP2;3  |
OsPIP2;4  |
OsPIP2;5  |
OsPIP2;6  |
OsPIP2;8  |
OsPIP2;7  |
| TaPIP1-2a | RTKS------- |
| OsPIP1;1 | KSRS------- |
| OsPIP1;2 | KSRS------- |
| OsPIP1;3 | KSRD------- |
| TaPIP2-9a | KALGSSFRNSR |
| OsPIP2;1 | KALGSSFRENA |
| OsPIP2;2 | K-LGSSFRNSA |
| OsPIP2;3 | K-LGSYRSNSA |
| OsPIP2;4 | G-YGSFRNSA |
| OsPIP2;5 | G-YGSFRNSA |
| OsPIP2;6 | NCSGSSGRGR |
| OsPIP2;8 | KAFASSYRS |
| OsPIP2;7 | KALSSFRSTS |

**ClustalW alignment of wheat TaPIP2-9a cds and TaPIP1-2a cds putative protein sequences with Rice PIP sequences.** Thick black lines above the alignment show the putative bilayer-spanning domains (TMH 1-6) and the connecting loops A – E are denoted as (LA, LB, LC, LD and LE). The amino acid residues forming four ar/R (arginine aromatic) selectivity filter are shown in ash and the amino acid residues shaded in black indicates the putative mercury sensitive sites (Guo et al., 2006a). The position of NPA motifs are boxed with rectangles in the Loop B and E, and the putative phosphorylation sites of PIP2 (Prak et al., 2008) are highlighted in blue. Green highlighted residues in PIP1 and orange highlighted residues in PIP2 sequences are indicative of PIP1 and PIP2 proteins of rice and wheat. Blue shading: position of AEF motif conserved in many MIPs (Zardoya and Villalba, 2001, Heymann and Engel, 2000).

Alignment created in ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and displayed in Bioedit (http://www.mbio.ncsu.edu/bioedit/bioedit.html).
Appendix 4 Accession numbers of deduced amino acid sequence of *H. vulgare*, *O. sativa*, *A. thaliana*, *T. aestivum* and *T. turgidum*.

<table>
<thead>
<tr>
<th>Intron 2</th>
<th>Species</th>
<th>Accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hordeum Vulgare</strong></td>
<td>HvPIP1-1</td>
<td>BAF41978</td>
</tr>
<tr>
<td></td>
<td>HvPIP1-2</td>
<td>BAF33067</td>
</tr>
<tr>
<td></td>
<td>HvPIP1-4</td>
<td>BAF33068</td>
</tr>
<tr>
<td></td>
<td>HvPIP2-1</td>
<td>BAE02729</td>
</tr>
<tr>
<td></td>
<td>HvPIP2-2</td>
<td>BAG06230</td>
</tr>
<tr>
<td></td>
<td>HvPIP2-3</td>
<td>BAF33069</td>
</tr>
<tr>
<td></td>
<td>HvPIP2-4</td>
<td>BAE06148</td>
</tr>
<tr>
<td></td>
<td>HvPIP2-5</td>
<td>BAG06231</td>
</tr>
<tr>
<td></td>
<td>HvPIP2-6</td>
<td>ADG03686</td>
</tr>
<tr>
<td></td>
<td>HvPIP2-7</td>
<td>ADW85675</td>
</tr>
<tr>
<td><strong>Triticum turgidum</strong></td>
<td>TaAQP2-1</td>
<td>ABW34454</td>
</tr>
<tr>
<td></td>
<td>TaAQP1-1</td>
<td>ABW34453</td>
</tr>
<tr>
<td><strong>Triticum aestivum</strong></td>
<td>TaPIP1AA</td>
<td>AAF61463</td>
</tr>
<tr>
<td></td>
<td>TaPIP2AA</td>
<td>AAF61465</td>
</tr>
<tr>
<td></td>
<td>TaPIP3AA</td>
<td>AAF61465</td>
</tr>
<tr>
<td></td>
<td>TaPIP1GG</td>
<td>AAM00368</td>
</tr>
<tr>
<td></td>
<td>TaPIP2GG</td>
<td>AAM00369</td>
</tr>
<tr>
<td></td>
<td>TaAQP1 (PIP1)</td>
<td>ABI96812</td>
</tr>
<tr>
<td></td>
<td>TaAQP2 (PIP2)</td>
<td>ABI96813</td>
</tr>
<tr>
<td></td>
<td>TaAQP3 (PIP2)</td>
<td>ABI96814</td>
</tr>
<tr>
<td></td>
<td>TaAQP4 (PIP2)</td>
<td>ABI96815</td>
</tr>
<tr>
<td><strong>Oryza sativa</strong></td>
<td>OsPIP1-1</td>
<td>Q6EU94</td>
</tr>
<tr>
<td></td>
<td>OsPIP1-2</td>
<td>Q7XSQ9</td>
</tr>
<tr>
<td></td>
<td>OsPIP1-3</td>
<td>Q9SX8</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-1</td>
<td>Q8HSN9</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-2</td>
<td>Q6K215</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-3</td>
<td>Q7XUA6</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-4</td>
<td>AAP54303</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-5</td>
<td>Q8GR8</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-6</td>
<td>Q7XLR1</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-7</td>
<td>Q651D5</td>
</tr>
<tr>
<td></td>
<td>OsPIP2-8</td>
<td>Q7Y1E6</td>
</tr>
<tr>
<td><strong>A. thaliana</strong></td>
<td>AtPIP1-1</td>
<td>CAB71073</td>
</tr>
<tr>
<td></td>
<td>AtPIP1-2</td>
<td>AAC28529</td>
</tr>
<tr>
<td></td>
<td>AtPIP1-3</td>
<td>AAF81320</td>
</tr>
<tr>
<td></td>
<td>AtPIP1-4</td>
<td>AAF02782</td>
</tr>
<tr>
<td></td>
<td>AtPIP1-5</td>
<td>CAA20461</td>
</tr>
<tr>
<td></td>
<td>AtPIP2-1</td>
<td>CAB67649</td>
</tr>
</tbody>
</table>

The table represents the Protein IDs of PIPs from barley (*Hordeum Vulgare*), wheat (*Triticum turgidum* and *Triticum aestivum*), Rice (*Oryza sativa*) and Arabidopsis (*A. thaliana*) used to construct phylogenetic tree.
Appendix 5 Expression pattern of Rice *OsPIP1-2* putative expressed protein.

The plot shows the expression of given genes during the developmental stages throughout the life cycle of an organism. A greater number of transcripts are expressed in the seedling stage (57 transcripts) compared to the other stages.
Appendix 6 (a) Alignment of predicted amino acid sequence of *TaPIP2-9a* cds obtained from root cDNA with all the known complete cds of wheat AQPs in the database.

The alignment of predicted amino acid sequence of the *TaPIP2-9a* and alternate splice variant gene obtained from this study with the known full length wheat PIP2 ESTs.

ABW34454 – *TaAQP2-1* (PIP2), ABI96815 – *TaAQP4* (PIP2), ABI96814 – *TaAQP3* (PIP2), AF366565 – *TaPIP2*, ABI96813 – *TaAQP2* (PIP2), AF139815 – *TaPIP2*. Predicted transmembrane helices are represented by black lines on top of the residues, Loops A, B, C, D and E are also represented on top of the residues.
Alignment of predicted amino acid sequence of \textit{TaPIP1-2a} \textit{cds} obtained from root cDNA with all the known complete \textit{cds} of wheat AQPs in the database.

The alignment of predicted amino acid sequence of the \textit{TaPIP1-2a} gene obtained from this study with the known full length wheat AQP ESTs.

The accession numbers are ABI96812 – TaAQP1 (PIP1), ABW34453 - TtAQP1.1 (PIP1), AF139814 (AAF61463) – TaPIP1 and AF366564 (AAM00368) – TaPIP1. NPA motifs are highlighted in ash.

Alignment created in ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and displayed in Bioedit (http://www.mbio.ncsu.edu/bioedit/bioedit.html).
Appendix 7

Uninfected insect cells before and after 72hrs incubation.

(a) and (c) Represents the uninfected insect cells incubated for 72hrs after and before processing with ImageJ, cells were densely populated and many dead cells were observed (cells with discontinuous outline in fig (a) and corresponding cells were very light in the fig (c)). Figure (b) represents the uninfected insect cells incubated at room temperature for 20minutes for the cells to settle down, cells look healthier and dead cells were not observed and corresponding unprocessed cells were shown in fig (d).
Appendices

Appendix 8

1. Select **File ➤ Open (ctrl+O)** from the menu and submenu of the ImageJ information box.

2. Choose the appropriate picture to be measured and double click on it.

3. Select **Image ➤ Adjust ➤ Brightness and contrast** from the side drop down menu.

4. Small **B&C** tool box appears and **auto** brightness/Contrast was selected.

5. Select **Process ➤ Smooth** from submenu.
Appendices

6. Select **Process ➤ Binary ➤ Make Binary** (Image will look like black holes against white background) from submenu.

7. Select **Process ➤ Binary ➤ Fill Holes** from submenu.

8. All the images (from 0.5sec to 10min, after adding sucrose and water) were processed and converted to black on white background.

9. Manually all the images (from 0.5sec to 10min) were compared to select the cells to be measured.

10. Maximum number of good looking cells (throughout 0.5sec to 10min) was selected to measure the surface area.

11. Select **Tracing** tool (shown in a black circle) from the ImageJ tool bar and click on the insect cell to be measured (one at a time), yellow circle around the cell.

    **measuring cells individually**

12. Select **Analyse ➤ Measure** from the submenu.
13. Result page appears for the cell that has been measured with values Area in pixel$^2$
(Note: do not close the result page until the entire cells in the image has been measured)

![](image1)

(b) Trial 1  (a) Trial 2

Trial (a) and (b) represents the difference between elliptical tool and tracing tool used in trial 1 and 2 (look at the result table).

14. Save the results from Results page, **File ➤ save as ➤ enter name ➤ enter** (will be saved in Excel sheet automatically).

15. All the cells from each image (0.5min, 1min, till 10min) in total 20 images were averaged.

16. Results from all the time points were imported into new excel sheet for easy comparison of the cells size from 0.5min to 10min.

17. All the averages were graphed in Excel sheet.