6. SUMMARY AND CONCLUSIONS

With increase in the number of completed genomes being published, determining protein function is one of the most important problems of this post-genomic era. In this study, we integrated knowledge based approach with the available standalone as well as online softwares for database searching, homology matching and molecular simulations to get an insight, albeit theoretically, into the behaviour of osmotin as cell signalling molecule. The molecular mechanism by which osmotin modulates the cell signalling cascade especially for proline accumulation and biosynthesis, leading to the adaptations by plants against stress, would be of great importance to the plant scientists. The following are the salient findings of the present study:

1. For osmotin to act as a gene regulator (transcription factor), it must contain a DNA binding motif. No DNA-binding motif was detected by DNA-binding motif prediction softwares; 2-zip server, DBD transcription factor prediction database, GYM 2.0, HTH motif prediction, and PredictDNAHth. Alignment results of osmotin protein with the protein sequence from DATF showed the homology in the range of 0-20%, suggesting that it does not contain a DNA binding motif. Further to find the DNA-binding domain, the superimposition of osmotin 3D structure was done on modelled Arabidopsis transcription factors using Superpose program that also suggested absence of the same. Thus, the overall results of transcription factor prediction softwares and alignment of osmotin protein with the protein sequences and structures of transcription factors from DATF suggest that osmotin do not regulate the cell signalling cascade as transcription factor.

2. The analysis of available plant genome sequences and existing annotations of the 11 plants (Arabidopsis lyrata, Arabidopsis thaliana, Carica papaya, Medicago truncatula, Glycine max, Cucumis sativus, Populus trichocarpa, Vitis vinifera, Oryza sativa, Zea mays and Sorghum bicolor) that were taken for this study, have shown the presence of osmotin and osmotin/ thaumatin like proteins in these plants. These osmotin/ thaumatin like proteins contain osmotin, either alone or in combination of 5 different partner domains; protein kinase-like (PK-like), bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin, mitochondrial carrier, Class II aaRS ABD-related and ARM repeats. Thus, the study showed that osmotin is ubiquitously present, especially in the genomes
taken in this study. It is a member of large superfamily of structurally similar proteins comprising of osmotin and osmotin like proteins. Therefore, apart from osmoregulation, members of this family may be involved in diverse biological roles including cellular signal transduction.

3. Due to the availability of its complete genome sequence, *Arabidopsis thaliana* genome was parsed for the presence of osmotin like protein using tobacco osmotin. The analysis showed 31 proteins belonging to Osmotin/thaumatin-like protein superfamily. Among these proteins, osmotin was present either alone or in combination with either of the two partner domains, Protein kinase-like (PK-like) (56112) and Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin (47699). The results obtained from Target P server for 31 members of *A. thaliana* Osmotin/ thaumatin like protein superfamily suggest that all, except proteins encoded by ten genes; AT4G18250.1, AT1G20030.1, AT1G73620.1, AT1G77700.1, AT4G36000.1, AT4G38660.1, AT4G38660.2, AT4G38670.1, AT4G38670.2, AT4G38670.3, are secretory in nature. The secretory nature and multiple locations targeting of osmotin and OLPs are in agreement to their multi functional role in plants exposed to biotic and abiotic stresses. All these factors indicate that osmotin is a member of large superfamily of structurally similar proteins in the *Arabidopsis* also.

4. Data mining showed that osmotin shares a common β-lectin fold with human adiponectin, a cell signalling molecule in oxidative stress tolerance, and exhibits it effect through a common receptor (ADIPOR1) in human myocytes. The BlastP of *Arabidopsis* genome with ADIPOR1 amino acid sequence showed the presence of hepta helical protein (HHP) family as receptor homologue for osmotin. It comprises of five members; HHP1, HHP2, HHP3, HHP4 and HHP5. HHP1 has 33 % homology to ADIPOR1 and share similar lengths of helices and loops between the helices.

5. Human ADIPOR1 and *Arabidopsis* HHP1 were modelled using fold-recognition and comparative modelling method. In absence of clear structural homolog, template was searched using alignment scores for secondary structure alignment for ADIPOR1 and HHP1, individually with the proteins containing the 7 helix trans-membrane topology, from OPM database (http://opm.phar.umich.edu/) using SSEA program along with the manual secondary structure alignment for lengths of helices and loops between the helices in the all the proteins containing single subunit. The results showed that the
structure of eubacteria \textit{(Anabena nostoc sp.)} sensory rhodopsin (PDB code: 1XIO) has similar lengths of helices and loops between the helices with respect to ADIPOR1 as well as HHP1. Therefore, this protein was used as the template for predicting the three-dimensional structure of ADIPOR1 by comparative modeling strategy.

6. Modelled structure of ADIPOR1 was obtained from Discover studio 2.0. The Ramachandran plot for the model showed 92.5\% residues in most favoured regions. The positioning of secondary structural elements, generated from PDBsum, showed that ADIPOR1 had seven helices, five continuous and two with breaks. Similarly, the modeled structure of HHP1, obtained from Discover studio 2.0, also showed seven helices, five continuous and two with breaks.

7. The binding of ligand to the receptor molecule induce structural changes in the receptor. To study the interaction between osmotin and HHP1, osmotin-HHP1 complex was modeled using Discover studio 2.0. Upon interaction, osmotin exposes three basic, two acidic, two aromatic and an uncharged hydrophilic residues, whereas the HHP1 exposes a basic, three uncharged hydrophilic, one aromatic and one cysteine residues. In particular, the amino acids located in \(\beta\)-strands, helices and loop in osmotin are exposed for interaction with HHP1 receptor upon binding. In turn HHP1 interacts by all the three extracellular loops as well as with its free C-terminal end. To analyze the interaction of ATOSM34 and HHP1, the ATOSM34-HHP1 complex was modeled using Discover studio 2.0. The results showed that upon interaction, ATOSM34 exposes one basic residue, an acidic and three uncharged hydrophilic residues where as the HHP1 exposes one basic residue, three uncharged hydrophilic and cysteine residue. In particular, the ATOSM34 exposes the amino acids located in \(\beta\)-strands, helices and loop, while HHP1 interacts by all the three extracellular loops as well as with its free C-terminal end. The interaction results of HHP1 with both osmotin and ATOSM34 show that all the three loops of HHP1 are involved in the interactions. The predominant interaction in osmotin-HHP1 and OSM34-HHP1 ligand receptor complexes are of electrostatic in nature.

8. To study the cell signaling pathway for osmotin induced proline biosynthesis and accumulation in plants further, submission of N terminal amino acid sequence of HHP1 to Group-based Phosphorylation Scoring Method (GPS) was done. Results showed the presence of phosphorylation sites of G-protein-coupled receptor kinases (GRKs) on the N-terminus HHP1 protein. The GRKs phosphorylation sites on the N-terminus HHP1
protein would initiate the binding of β-arrestin like proteins leading to the activation of ERK mediated phosphorylation of enzymes and transcription factors involved in proline biosynthesis and accumulation.

9. To dissect the pathway further, we determined the enzymes and transcription factors of proline biosynthesis, catabolism and transport. A set of seven genes (Delta1-pyrroline-5-carboxylate synthetase 1, Pyrroline-5-carboxylate reductase, Proline oxidase, Putative / osmotic stress-responsive proline dehydrogenase, Delta-pyrroline-5-carboxylate dehydrogenase, Proline transporter 1, Proline transporter 2, Proline transporter3) associated with proline biosynthesis, catabolism and transport were identified. The presence of O2 and OCSBF-1 consensus binding sites, occurring in combination, within the promoters of genes involved in proline biosynthesis and accumulation have been shown using composite model of ExPlain™ Plant. These findings of our study suggest a role for bZIP family proteins in the regulation of expression of this gene set.

10. To understand further downstream pathway, ERK phosphorylation sites on amino acid sequences of the enzymes and transcription factors of proline biosynthesis and accumulation were determined. Upon analyses, P5CS1, the enzyme of proline synthesis pathway showed two sites for ERK phosphorylation both followed by well defined D domain. PROD and P5CDH, the enzymes of proline catabolism pathway also showed two sites for ERK kinase phosphorylation. In PROD, both ERK phosphorylation sites were followed by well defined D domain. In P5CDH, however, one ERK phosphorylation site was followed by D domain, while the other by FXFP domain. Similar analyses were performed for the transcription factors, AtbZIP10, AtbZIP25, AtbZIP53, AtbZIP2, MYBAS1 and GT-1. Among these, only AtbZIP25 showed single site as phosphoacceptor motif for ERK.

11. Finally, on the basis of evidences obtained in our study, we have concluded that apart from antifungal activity, osmotin and its subfamily proteins at molecular level act as cell signalling molecule(s) through a signalling cascade employing HHP1 receptor, β-arrestins like protein and ERK MAP kinases. Using this cascade, they modulate proline metabolism in plants by regulating gene expressions of proline biosynthetic and proline catabolic enzymes through transcription factor (AtbZIP25). At protein level these signalling molecule enhance synthesis and accumulation of proline in plants by increasing the activity of proline biosynthetic enzyme (P5CS1) while down regulating the
activities of proline catabolic enzymes (PROD and P5CDH) via phosphorylation. The sequence of events that followed in the signalling pathway for modulation of proline biosynthesis and accumulation by osmotin protein are presented by a model proposed by us based on the evidences collected in the present study:

Under the stress conditions, the expression of osmotin subfamily genes may increase in plants leading to both intracellular and extracellular accumulation of osmotin proteins. Intracellular osmotin plays a role in osmotic adjustment of the cell, whereas the extracellular osmotin may bind to the plasma-membrane receptor HHP1, a homologue of ADIPOR1. Thereafter, GRK like kinases may phosphorylate HHP1 leading to the binding of β-arrestins like protein. This may provide a scaffold for binding and activation of ERK that in-turn may activate the transcription factor, AtbZIP25. It may either increase transcription of genes for proline biosynthetic enzymes by binding to their promoter/enhancer region and /or repress transcription of genes for proline degradation enzymes by binding to their promoter/silencer region. ERK may also enhance biosynthesis and accumulation of proline by activating proline biosynthetic enzyme (P5CS1) or deactivating the enzymes of proline catabolism (PROD and P5CDH) via phosphorylation (Figure 32).

Thus our method is the first of its kind to figure out, albeit theoretically, potential proteins enacting in this signalling pathway. Whether this intricate system formed by combination of signal, membrane receptor and molecular scaffolds is actually utilized by cells to regulate the proline biosynthesis, has to be determined by the experimental methods. The reverse genetic approaches (random or specific gene disruption, over-expression of strong activators or repressors or their inducible counterparts) in combination with microarray technology could be used to detect the target genes and the identification of their interacting partners. Such studies with the metabolomic analyses should pave the way towards a better understanding of the functional diversification of plant transcription factors and regulatory molecules under stress conditions.