Molecular Characterisation of Inositol Monophosphatase like (IMP-L) Proteins in Chickpea

A Thesis Submitted
in Partial Fulfillment of the Requirement for the Degree of

Master of Technology

In

Industrial Biotechnology

by

Bidisha Bhowal

Registration No.- DTU/M.TECH/2K15/57

Department of Biotechnology

Delhi Technological University

Under the supervision of

Dr. Prakash Chandra
Assistant Professor
Department of Biotechnology
Delhi Technological University

Prof. D Kumar
Head of Department
Department of Biotechnology
Delhi Technological University

Dr. Saurabh C. Saxena
Assistant Professor
Department of Biotechnology
Delhi Technological University

DECLARATION

I, Bidisha Bhowal, M.Tech Industrial Biotechnology, bearing **Registration No. DTU/M.TECH/2K15/57**, Department of Biotechnology, Delhi Technological University, Delhi declare that my project work titled "***Molecular characterisation of Inositol monophosphatase like (IMPL) proteins in Chickpea***", is original and no part of this work report has been submitted for any other degree or diploma. All the given information and works are true to my sense and knowledge.

Date:

Place: New Delhi

Bidisha Bhowal

CERTIFICATE

This is to certify that the major project entitled *"**Molecular characterisation of Inositol monophosphatase like proteins in Chickpea**"* is the bonafide work of **Ms. Bidisha Bhowal**, a M.Tech (Industrial Biotechnology) student from Delhi Technological University, Delhi, is a record of her own work carried out under our supervision and guidance. The information and data enclosed in this thesis is original and has not been submitted elsewhere for honouring of any other degree.

Dr. Prakash Chandra
(Supervisor)
Assistant Professor
Department of Biotechnology
Delhi Technological University

Prof. D Kumar
(Co-Supervisor)
Head of Department
Department of Biotechnology
Delhi Technological University

ACKNOWLEDGEMENT

I would like to take this opportunity to express my indebtedness to Prof D. Kumar, Head of the Department, Department of Biotechnology, Delhi Technological University for allowing me to pursue my project at NIPGR, New Delhi.

I would also like to express my sincere gratitude to Dr. Prakash Chandra, Assistant Professor and, Dr. Saurabh Saxena, Assistant Professor Department of Biotechnology, DTU, Delhi; for introducing the present topic and for their inspiring guidance, constructive and valuable suggestion throughout this work.

I am heartily thankful for their guidance and support during my project. Their able knowledge and expert supervision with unswerving patience fathered my work at every stage, for without their warm affection and encouragement, the fulfilment of the task would have been very difficult.

I wish to express my deep sense of gratitude and indebtedness to Dr. Manoj Majee, Scientist, National Institute of Plant Genome Research for his highly valuable guidance and constant inspiration throughout the work. I would like to thank Director, NIPGR for providing the institutional infrastructure and numerous other facilities for carrying out my project

I am genuinely grateful to all my seniors in the lab, Dr. Pooja Verma, Dr. Shweta Roy, Dr. Prafful Salvi, Dr. Venketeshwara Rao, Shraboni Ma'am, Abhijit Sir for their suggestions and moral support during my work. I would like to express special gratitude and respect to Nitin Sir for his constant mentoring and guidance throughout the project. This work would not have been possible without his encouragement and support.

I am also thankful to Nandan bhaiya and Rahul bhaiya for their help during the project.

A special note of thanks to my family and friends for keeping me high spirited throughout.

CONTENTS

ABBREVIATIONS

IMPL-Inositol Monophosphatase like protein

mg- milligram

LB media- Luria Bertini Miller media

MS media-Murashige and Skoog media

BSA-Bovine Serum Albumin

APS-Ammonium persulphate

TEMED-N,N,N,N-Tetraethylenemethyldiamine

β-ME- β-mercaptoethanol

IPTG-Isopropyl- β,D thiogalactoside

µg-microgram

µL-microlitre

NaCl-Sodium chloride

% - percentage

CBB R250 -Coomassie Brilliant Blue R 250

TAE-Tris-Acetate EDTA

EDTA-Ethylene Diamine Tetra Acetic acid

Kan- Kanamycin

Amp-Ampicillin

Rif-Rifampicin

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Inositol holds great importance for the physiological well being of organisms. In eukaryotes especially, the intracellular pools of inositol are essential metabolites as well as labile messengers that regulate cellular physiology while travelling within and between cells. Free inositol is synthesised by either the *de novo* biosynthetic pathway or recycling of inositol phosphates by de-phosphorylation. It is a multifaceted compound that generates myriad of derivatives upon phosphorylation. The key enzyme needed for the *myo*-inositol biosynthesis functionally, a hydrolase, is known as *myo*-inositol monophosphatase (IMP). In addition to IMP, two other genes have been hypothesised to have IMP activity that has resulted in their classification as IMP-like (IMPL) proteins. The IMP together with IMPL constitutes the inositol monophosphatase gene family. Inositol monophosphatase (IMP) gene has previously been characterised in chickpea. In this project, we aim to isolate and clone the gene encoding *IMPL1* and *IMPL2* and its corresponding cDNA from chickpea. Also, we intend to study the bioactivity of the proteins using bacterially over-expressed and purified protein fractions. Further, we also aim to functionally characterise CaIMPL1 and CaIMPL2.

1. INTRODUCTION

Myo-inositol incorporates itself into many vital cellular compounds such as the ones involved in the production of lipid phosphatidylinositol phosphate signalling molecules, thereby playing a pivotal role in the growth and development of organisms. This holds true especially in plant biology where molecular entities containing or utilizing *myo*-inositol such as glycerophosphoinositide membrane anchors, cell wall pectic, non cellulosic polysaccharides, and ascorbic acid are synthesised.

Myo-Inositol monophosphatase (IMP) (3.1.3.25) which is functionally a hydrolase, is a crucial enzyme in the metabolic pathway of *myo*-inositol. It is basically involved in the de-phosphorylation of Inositol-1-phosphate but is also known to catalyse the de-phosphorylation of other inositol phosphate compounds like the breakdown products of phosphoinositides (Gillaspy et al., 1995; Quintero et al., 1996; Loewus and Murthy, 2000) for the maintenance of the cellular inositol pool. Thus, this enzyme can be contemplated to be a crucial and potent regulatory point for all pathways which are involved in the usage of free inositol.

The IMP proteins are members of the metal-dependent phosphatases' super-family having specificity towards wide range of substrates. Till date, three probable IMP encoding sequences have been identified in the Arabidopsis genome: *VTC4 (At3g02870)*, *Myo*-inositol monophosphatase-like1 (*IMPL1; At1g31190), and IMPL2 (At4g39120)*; (Torabinejad et al., 2009). *VTC4* has been recently reported to be a bi-functional enzyme catalysing the breakdown of the L-Gal 1-P in ascorbate biosynthesis in Arabidopsis (Laing et al., 2004; Conklin et al., 2006), and also is able to catalyze the de-phosphorylation of D-Ins 1-P and D-Ins 3-P to myo-inositol *in vitro*. Analyses of *imp* mutants revealed a 30% decrease in *myo*-inositol content suggesting the redundant enzyme functioning in synthesis and metabolism of inositol, like the two IMPL proteins which have been shown to have *in vitro* IMP activity (Torabinejad et al., 2009)

These proteins are known as IMP-like (IMPL), because they are closer in sequence identity to the prokaryotic IMPs, namely the SuhB (Matsuhisa et at., 1995; Chen and Roberts, 2000) and CysQ (Neuwald et al., 1992; Peng and Verma, 1995) proteins. There have been several reports suggesting the presence of IMP-like proteins even in plants.

The inositol monophosphatase gene family has been previously characterised in Arabidopsis to study the functioning of the gene family in the synthesis and metabolism of inositol. The study suggests that IMPL2 functions in the histidine biosynthetic pathway, while IMPL1 catalyse the hydrolysis of inositol - and galactose-phosphates in the plant cell.

Inositol metabolism has been previously reported to trigger a positive response to drought stress in chickpea (Boominathan et al., 2004). Chickpea is subjected to different abiotic stresses during its life cycle owing to its growth in arid and semi-arid regions. As a result, chickpea harbours many stress resistant genes particularly for abiotic stresses such as drought, cold, and salinity. The role and regulation of one such gene encoding inositol monophosphatase has been studied. The study reveals that CaIMP improves seed germination and seedling growth linking various metabolic pathways, especially when subjected to stress conditions (Saxena et al., 2013).

Since IMPLs are hypothesised to contain IMP activity, we intend to study the detailed enzymatic properties of bacterially over-expressed and purified CaIMPL1 and CaIMPL2.

OBJECTIVES

In accordance, the project has been designed with the following objectives:

- Isolation and cloning of *CaIMPL1* and *CaIMPL2* cDNA
- Biochemical Characterisation of CaIMPL1 and CaIMPL2
- Functional Characterisation of CaIMPL1 and CaIMPL2

2. REVIEW OF LITERATURE

2.1 Chickpea (*Cicer arietinum* L.)

Chickpea (*Cicer arietinum*) (2*n* = 16) is the world's second most widely grown legume. Its cultivation is particularly important to food security in the developing world. Owing to its capacity for symbiotic nitrogen fixation, chickpea seeds are a primary source of human dietary protein (Jukanti A.K et al., 2012). Chickpea belongs to the Papilionoid subfamily of legumes, a clade that contains essentially all of the important legume crops. Within this subfamily, chickpea is most closely related to crops such as alfalfa (*Medicago sativa*), clover (*Trifolium* spp.), pea (*Pisum sativum*), lentil (*Lens culinaris)*, and the model legumes barrel medic (*Medicago truncatula*) and *Lotus japonicus*. Soybean (*Glycine max*). Originating in south east Turkey and Syria, chickpea is considered to be the founder crops of modern agriculture (Maesan et al., 1987, Zohary et al., 1983)

In many parts of the world, chickpea is cultivated in semi-arid environments and on soils of poor agricultural quality, which, combined with its susceptibility to drought and debilitating fungal diseases, have restricted yields to <1 ton/ha, which is considerably below the theoretical potential. Genetic improvement, either by traditional or molecular methods, has been hampered by the limited genomic resources coupled with narrow genetic diversity in the elite gene pool (Varshney et al., 2010)

2.1.1 Classification

Kingdom: Plantae

Division: Magnoliophyta

Class: Magnoliopsida

Order: Fabales

Family: Fabaceae (Leguminosae)

Subfamily: Faboideae(Papilionaceae)

Genus: *Cicer*

Species: *arietinum*

2.1.2 Response to stress

The susceptibility of chickpea to dehydration severely reduces the yield and its productivity has remained historically low. Plant cell wall or extracellular matrix (ECM) is the first compartment that senses the stress signals, transmits them to the cell interior, and eventually influences the cell fate decision. It is now well established that plants being sessile have evolved many adaptations to counteract dehydration.

These adaptations are classified into four categories:
1.Dehydration avoidance (developmental and physiological traits)
2.Dehydration tolerance (physiological and biochemical adaptations)
3.Dehydration escape and dehydration recovery

A few characteristics such as osmotic adjustment (OA) and cell membrane stability are recognized as effective components of dehydration tolerance in many crops. These are expressed in terms of relative water content (RWC) of the plant, accumulation of compatible solutes like proline, and increased permeability of ions and electrolytes. Further, the status of photosynthetic machinery has been considered as an ideal index to monitor the health and vitality of plants during dehydration.

Inositol metabolism is known to play an important role in chickpea, especially in response to drought stress (Boominathan et al., 2004). Chickpea faces various environmental stresses during its life cycle and has therefore been considered as a rich source of tolerant genes for a wide range of abiotic stresses such as drought, cold, and salinity. It has been the subject of major breeding programmes worldwide in concern to the various biotic and abiotic stresses affecting this crop.

2.2 Inositol and its derivatives: evolution and functions

The inositols are the nine isomeric forms of cyclohexanol, with the empirical formula $C_6H_{12}O_6$, and formula weight of 180.16. It is a group of small and chemically very stable polar molecules that have versatile properties. *Myo*-inositol is the most used form in biology. The inositols constitute a subgroup of a broader class of compounds known as cyclitols-polyhydroxylated and non-reducing compounds in which the core structure is an all-carbon ring of six(or occasionally five or seven) carbons, in which at least three ring carbons have directly attached hydroxyl groups (Posternak T, et al.,1965).

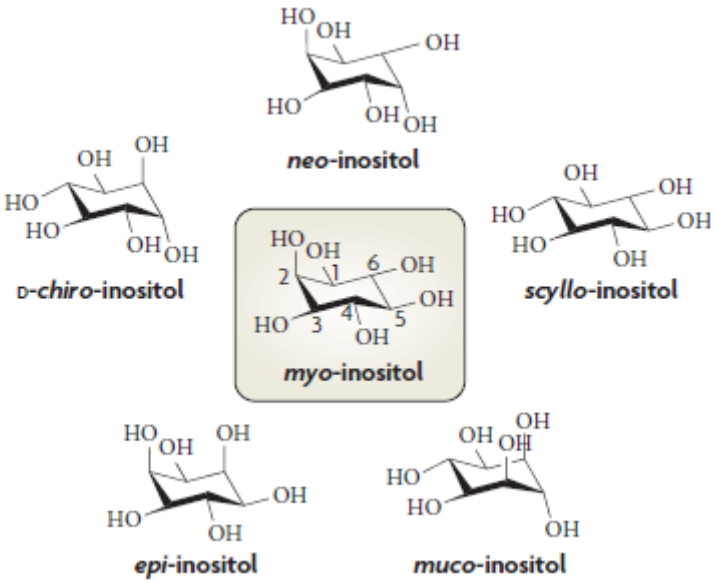


Figure A1: The diversity of inositols and their derivatives.

2.3 Inositol distribution in three kingdoms

The importance, distribution and usage of inositol derivatives varies across the three biological kingdoms. Archaea, Bacteria and Eukarya. Inositol and its derivatives are made and used by most archaea and are prevalent in eukaryotes, whereas only relatively few bacteria make use of them

Inositol and its derivatives are not only used by rather few bacteria, but also have very few known functions in the mitochondrial or chloroplast descendants of α-proteobacteria or cyanobacteria.

All eukaryotic cells use inositol always in their membrane phosphoinositides and usually for diverse other processes. Some cells can either harvest available inositol or make their own when none is available (Nunez L, et al., 2006). In yeast, this is achieved by transcriptional regulation of the synthesis of *myo*-inositol phosphate synthase by the environmental inositol concentration (Carman G.M., et al., 2004). Some cultured mammalian cells can make their own inositol but others require an exogenous supply (Eagle H. et al., 1957, Jackson M. et al., 1982).

2.4 Eukaryotic inositol usage

Inositol containing phospholipids are plentiful constituents of the membranes of many archaea and of all eukaryotes. Many eukaryotes use inositol-based stress protective cytosolic solutes, all probably need plasma membrane phosphatidyl-inositol-4,5-bisphosphate for endocytosis, exocytosis and sub plasmalemmal cytoskeleton integrity. All use phosphatidylinositol-3-phosphate and probably also phosphatidylinositol-3,5-bisphosphate to regulate membrane trafficking in secretory and endocytic pathways. Hydrolysis of phosphatidyl-inositol-4,5-bisphosphate to the second messengers inositol-1,4,5-triphosphate and sn-1,2-diacylglycerol by phosphoinositidase C is ubiquitious in eukaryotes and the liberated inositol-1,4,5-triphosphatealso serves at least sometimes as a precursor of phytic acid and pyrophosphates. Most of the functions of myo inositol lipids are brought about as a result of their interactions with diverse proteins that interact specifically with particular lipid

2.5 Nomenclature

Myo inositol is a *meso* compound with a plane of symmetry that rotates the structure about C2 and C5 as fixed positions. The remaining four carbon atoms consist of two prochiral pairs, C1=C3 and C4=C6. If the carbon ring is numbered clockwise, as shown by numbers inside the ring, assignment of single substituent on carbon 1 is 1L.Conversely, if the carbon ring is numbered counter-clockwise, as shown by numbers external to the ring, assignment is 1D.

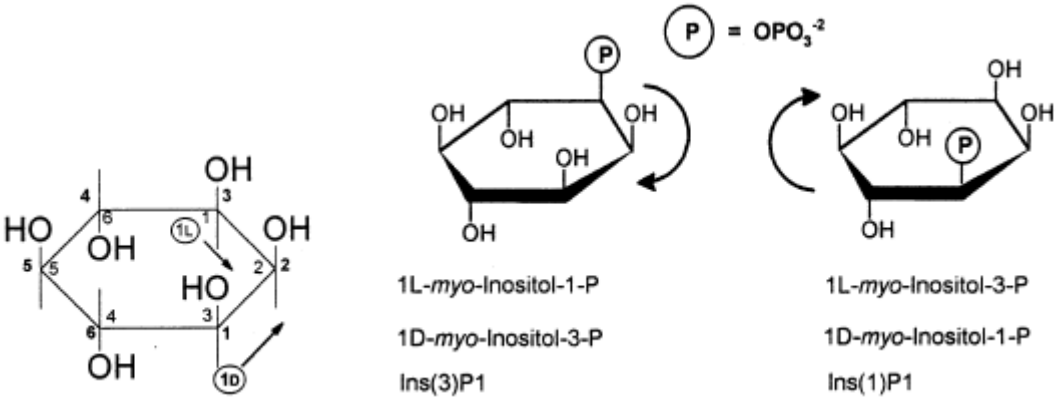


Figure A2: Convention for numbering substituents in *myo*-inositol

The upshot is a tentative agreement by the International Union of Biochemistry to relax rules of nomenclature so that 1L-MI-1-P,the product of myo inositol -1-phosphate synthase, may be designated 1L-MI-1-P,1D-MI-1-P,or simply Ins(3)P1 where the symbol Ins signifies *Myo* inositol with counter-clockwise numbering from 1D.

2.6 *Myo*-inositol Biosynthesis

Myo-inositol was first isolated from muscle extract by Scherer. In the 1930s, Dangschat and Posternak published its full chemical structure. Although Maquenne in 1887 identified the D-gluco configuration of *myo*-inositol but Fischer in 1945 gave the idea of an enzymatic conversion of D-Glucose-6-Phosphate to form Inositol-3-Phosphate.The experimental evidence for such conversion came in 1962 when radio labelled glucose was traced to myo inositol in parsley leaf and the reaction was shown as cyclization of the carbon chain of D-glucose to form myo inositol. Finally, it was revealed that this conversion involves a mechanism of internal oxidation-reduction reaction involving Glucose-6-phosphate and NAD+

The synthesis of *myo*-inositol consumes the central glycolytic metabolite Glucose-6-phosphate.

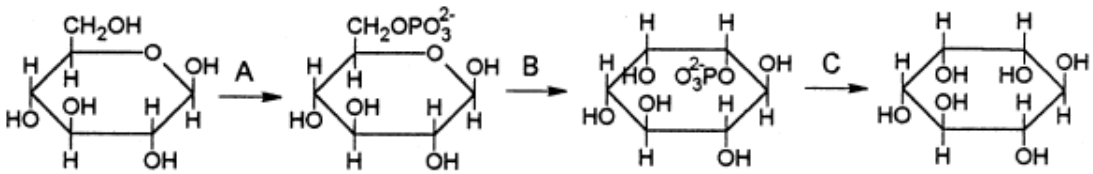


Figure A3: Conversion of D glucose to myo inositol

(A): Hexokinase (B): Inositol 3 phosphate synthase (C): myo inositol monophosphatase

The conversion of glucose-6-phosphate to *myo*-inositol-3-phosphate is the first committed step in the biosynthesis of *myo*-inositol (Loewus F.A *et al*., 1983). This cyclization step is irreversible and catalysed by the rate limiting enzyme *myo*-inositol-3-phosphate synthase [MIPS E.C 5.5.1.4]. In the last step, *myo*-inositol-1-phosphate is de-phosphorylated by specific Mg^{2+} dependent myo inositol phosphate phosphatase (IMP E.C 3.1.3.25) to produce free inositol.

This scheme constitutes the sole pathway of myo inositol biosynthesis in host of organisms ranging from fungi, algae cyanobacteria and occupies a central role in their cellular metabolism.

There exists very limited evidence that suggests that the other biological inositols may be made from myo inositol by simple inversion of configuration (epimerization) of one or two myo inositol hydroxyls (Hipps P.P, et al., 1973). Oxidation of inositol by *myo*-inositol oxygenase (MIOX) produces D-GlcA, which is possible entry point into ascorbate synthesis.

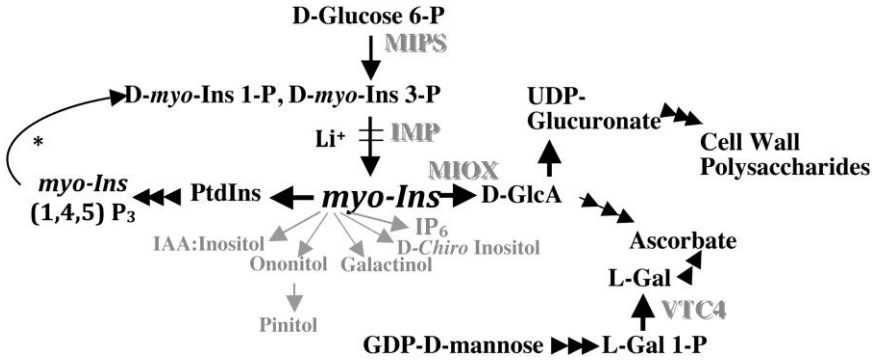


Figure A4: Myo-inositol synthesis and metabolism pathway.

Inositol is also the precursor for the synthesis of several compounds indicated in gray. The asterisk indicates inositol signalling pathway.

2.7 Role of *myo*-inositol in plant metabolism

The fact that *myo*-inositol plays a central role in growth and development in organisms is especially true in case of plant biology where molecular entities containing or utilising *myo*-inositol are involved in structure and function.

Metabolic processing of myo-inositol beyond biosynthesis produces other stereo forms of inositol and leads to a range of diverse functional roles.

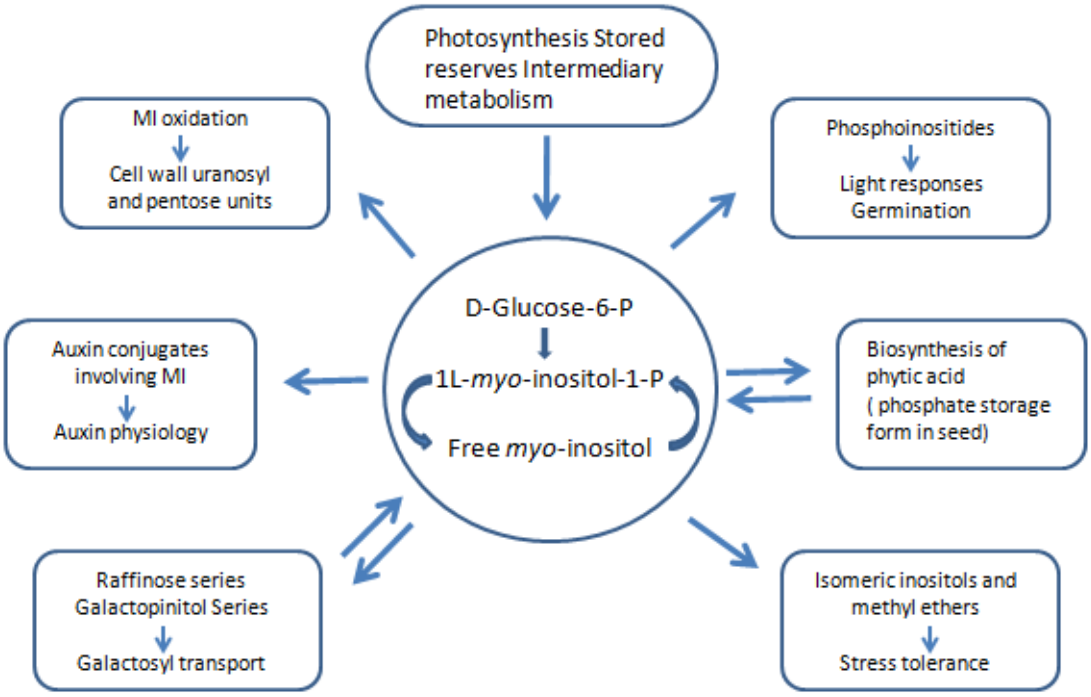


Figure A5: Functional roles of *myo*-inositol in plant metabolism

1. Cycling of 1L-MI-1-P and free MI by MI phosphatase and MI kinase (Gillaspy et al., 1995; Loewus et al., 1982)

2. Oxidation of free MI to D-glucuronic acid and further to uranosyl and pentanosyl units of pectin and hemicellµlose involved in cell wall biogenesis

3. Esterification of MI to form auxin esters and their glycosides (Slovin J.P et al., 1990)

4. Conjugation of free MI with D-galactose to form galactinol,the galactosyl donor for biosynthesis in the raffinose series of oligosaccharides (Horbowicz et al., 1994; Obendorf et al., 1997)

5. Isomerization and methylation of MI and other isomeric inositols to form ononitol, pinitol which participate in stress related responses (Horbowicz et al., 1994; Obendorf et al., 1997; Bohnert et al., 1996; Petrbauer et al., 1998)

6. Biosynthesis of phytic acid and phytic acid pyrophosphates (Murthy et al., 1996; Cosgrove et al., 1980; Brearley et al., 1996)

The role of inositols in biotic stresses has not been taken much seriously and there are very few reports that establish any connection between them. But the role of inositol derivatives in signalling cascade is well defined and an overlap between biotic and abiotic stress signalling cannot be ruled out. Transgenic potato plants carrying antisense constructs for MIPS were found to be less resistant to avirulent pathogen potato virus Y alongwith virulent strain of TMV. These plants were shown as having low levels of inositol hexakisphosphate (Murphy et al., 2008)

Inositol phosphates and their turnover products have been implicated to play important roles in stress signalling in eukaryotic cells. The stress induced increased accumulation of inositol has been reported in a few plants including chickpea. It has been shown that IMP activity is distributed in all organs in Chickpea and was noticeably enhanced during environmental stresses. Moreover, *Arabidopsis* transgenic plants over-expressing *Ca*IMP exhibited improved tolerance to stress during seed germination and seedling growth (Saxena et al., 2013)

Chapter 3

Material and Methods

3.1 MATERIALS

3.1.1 General Chemicals Used

Agarose, Ammonium per sulphate, Acrylamide (N,N'-methylene) -Bis acrylamide, β-mercaptoethanol, Bovine serum albumin (BSA), Chloroform/Isoamylalcohol (24:1) Coomasie Brilliant Blue R250, Calcium chloride, D-Glucose-6-phosphate, Ethanol, Electrophoresis buffer, Isopropanol, IPTG, NaCl, Pre-stained protein markers, Polyethylene Glycol, RNase A, Sodium acetate, TEMED, were purchased from Sigma Aldrich.

3.1.2 Antibiotics

Kanamycin, Ampicillin, Rifampicin, Hygromycin

3.1.3 Microbial Strains, Vectors and Enzymes

All microbial strains, vectors and enzymes used in this project are mentioned in Appendix-B

3.1.4 Growth Conditions of Chickpea and Arabidopsis

After thoroughly washing and soaking chickpea seeds in distilled water, overnight, they were kept in dark at room temperature for germination, after which the seeds were grown in pots filled with agropeat and vermiculite in a 3:1 ratio, under controlled conditions of 20^0C and12^0C of day and night temperature respectively with 50% humidity.

Arabidopsis thaliana ecotype Columbia seeds were grown in 3:1 agropeat and vermiculite combination under controlled conditions of 16/8h light and dark cycles and light intensity of 100µM/m/2/s). Arabidopsis seeds were sterilised by washing with 70% ethanol for 2 minutes followed by 1.2% sodium hypochlorite solution and kept for 15 minutes with shaking. Further, the seeds were washed with autoclaved MilliQ water.

3.1.5 Buffers, Solutions and Media Composition

All buffers, solution and media composition used in this project have been mentioned in Appendix-C

3.1.6 Sequencing

Sequencing was done in the institute using ABI3730XL sequencer.

3.2 METHODOLOGY

3.2.1 Standard Methods

3.2.1.1 Isolation of Total RNA by Tri Reagent

1. Tissue samples were homogenised in the TRI reagent (1 ml per 50-100mg of tissue) and allowed to stand for 5 minutes at room temperature to which 0.2 ml of chloroform was added per ml of TRI-reagent.
2. The sample was shaken vigorously for 15 second and thereafter kept at room temperature for 2-15 min.
3. Thereafter, centrifuged at 12,000 g for 15 min at 4^0C
4. The aqueous phase was collected in a fresh 1.5 ml micro-centrifuge tube to which 0.5ml of iso-propanol was added per ml of the TRI reagent used in sample preparation. The sample was allowed to stand for 5-10 min at room temperature.
5. Centrifuged at 12,000g 10min at 4^0C
6. The RNA precipitate forms a pellet on the side and bottom of the tube.
7. 1ml of 75% ethanol per ml of TRI reagent was used to wash the RNA pellet followed by vortexing and centrifugation at 7500 g for 5min at 4^0C.
8. The RNA pellet was briefly dried for 5-10 min and then dissolved in 30µl of RNase free water repeatedly pipetting and stored at -20^0C

RNA Quantification

Isolated RNA samples were quantified using Nanodrop instrument. For this, 1µl of sample was loaded and A260/280 ratio was determined.

3.2.1.2 DNAse Treatment

Components	Volume
RNA 10ug	Upto 21.5µl
10X DNase I buffer	2.5µl
DNase I	1.0µl
RNA free water	Required volume
Total volume	25µl

Table 1: Components of DNAse treatment

1. The reaction mix was incubated 37^0C for 30 minutes.
2. 5µl of DNase inactivation reagent was added by vortexing the tube followed by incubation at room temperature for 2 minutes and thereafter, by centrifugation at 10000 rpm to pellet the DNAse reagent.

3.2.1.3 cDNA Synthesis

cDNA was synthesised using reverse transcriptase and oligo-dT anchor primer.

Components	Volume
5x iScript Reaction mix	4µl
iScript Reverse Transcriptase	1µl
RNA template (DNAse treated)	15µl
Total volume	20µl

Table 2: Components for cDNA synthesis

1. The reaction mix was prepared
2. PCR tubes were aliquoted and given a short spin.
3. Tubes were then put into thermocycler and settings were adjusted to

42^0C	30 minutes	1 cycle cDNA synthesis
85^0C	5 minutes	1 cycle cDNA inactivation

3.2.1.4 Gel Electrophoresis

1. 50ml of 0.8% agarose gel was prepared by adding 0.4g of agarose in 50ml 1X TAE buffer
2. The mix was heated until a clear solution was obtained.
3. The solution was then allowed to cool to about 55^0 C before pouring. EtBr was added at a concentration of 0.5µg/µl.
4. The solution was then poured into the casting tray with the well comb in place and allowed to solidify for 20 minutes at room temperature.
5. Once solidified, the gel was placed into the electrophoresis unit and filled with 1X TAE until the gel got covered.
6. 1µl of 10X DNA loading dye was added to each of the digested samples
7. Molecular weight ladder was carefully loaded into the first lane and the samples in the remaining wells of the gel.
8. The samples were then electrophoresed at 80-150V until the dye reaches approximately 75-80% of the way down the gel.

3.2.1.5 Elution of DNA Using Gel Extraction Kit

1. The gel slices containing the DNA fragments were excised and put into a 1.5ml tube and the weight recorded followed by addition of binding buffer in a 1:1 volume
2. The gel mixture was then incubated at 55^0C for 10 min or until the gel slice completely dissolved.
3. After complete dissolution of the gel slice, 200µl of isopropanol was added to the samples.
4. The samples were then applied to the column and centrifuged at 12,000 rpm for 1 min.
5. This was followed by adding 100µl of binding buffer after discarding the flow through. The sample was centrifuged at 12000rpm for 1 min.
6. 700µl of wash buffer was added after discarding the flow through, centrifuged at 12000 rpm for 1 min.
7. The columns were centrifuged at the same stopped for an additional 1 minute for the removal of excess wash buffer.
8. The columns were then placed in clean 1.5ml micro centrifuge tubes.
9. Elution of DNA was done by dispensing 35µl of sigma water at the centre of the columns followed by standing for 1 minute and thereafter centrifuged at 12000rpm for 1 minute.

3.2.1.6 Restriction Digestion and Ligation of DNA

DNA and plasmids were digested using restriction enzymes and their respective buffers at the recommended temperatures. Ligation of DNA insert with vector was carried out using T4 DNA ligase followed by overnight incubation at 20-22^{0}C.

3.2.1.7 Cloning in pJET1.2 Vector

The purified PCR products were then used for ligation into pJET1.2 vector

Component	Volume (µl)
2X reaction Buffer	5
PCR product	3
T4 DNA ligase	0.5
pJET 1.2 vector	0.5
Nuclease free water to make up volume upto	10

Table 3: Ligation mixture components

The ligation mixture was incubated at room temperature (22°C) for 30 min and thereafter used for transformation

3.2.1.8 Preparation of Competent Cell

1. 500µl of overnight grown culture was transferred to 25ml of Luria Bertini liquid culture medium and incubated at 37^{0}C for 2-3 hrs with vigorous agitation.
2. Whole cell volume was transferred to 50 ml of ice chilled tube and kept in ice for 10 minutes, the cell volume was then centrifuged at 4100rpm for 10 minutes at 4^{0}C
3. Pellet formed was resuspended by vortexing in 12.5ml (half of initial culture volume, x/2) of 0.1M ice cold $CaCl_2$ solution.
4. Centrifuged again at 4100rpm for 10 min at 4^{0}C.
5. Pellet was again resuspended in 12.5ml (x/2) of 0.1M ice cold CaCl2 solution and stored in ice for 20 min to develop the competent cell
6. Centrifuged at 4100rpm for 10 minutes at 4^{0}C
7. Pellet was resuspended in 2ml of 0.1M $CaCl_2$ solution and stored in ice for transformation
8. 100µl of competent cell was taken and 16% glycerol added and stored at -80^{0}C for further use.

3.2.1.9 Bacterial Transformation

1. 2µl of DNA was gently mixed with 100µl of competent cells and kept in ice for 30 minutes.
2. This was followed by transferring the cells to water bath at 42^0C for 90 second (Heat shock) and thereafter kept in ice for 5 minutes.
3. 80µl of LB broth was added to each of the tube, incubated in a shaker at 37^0C, 180 rpm for 1 hour 30 minutes (1hour in horizontal condition for plasmid DNA).
4. Centrifuged at 4100 rpm for 7 minutes (Not required in case of plasmid DNA)
5. Excess media was removed by discarding the supernatant saving 100µl to suspend cells. Whole cell volume was then spread onto LB+antibiotic plates

3.2.1.10 Colony PCR

1. Few colonies were selected and screened for the presence of insert by colony PCR according to the protocol described by Sambrook and Russell (2001).
2. Suspension of the colonies was prepared by dissolving a miniscule amount of inoculum into 10 µl of autoclaved water. To the colony suspension, buffer, DNA polymerase, dNTPs and respective primers were added. PCR was setup as follows

95^0C	5 min	1 cycle
95^0C	30sec	35 cycles
55^0C	1 min	
72^0C	1min 30sec	
4^0C	∞	Hold

Table 4: Reaction conditions for Colony PCR

3.2.1.11 Plasmid Isolation using pDNA Miniprep Kit

1. 1.5ml of overnight bacterial culture was centrifuged in a micro centrifuge tube at 13000rpm for 30 seconds.
2. The pellet was resuspended in 250µl of buffer AL1 containing RNaseA by vortexing or pipetting.
3. 250µl of buffer AL2 was added and mixed gently but thoroughly by inverting the tube 4-6 times.
4. Next 350µl of buffer AL3 was added and mixed immediately but gently by inverting the tube 4-6 times.
5. The sample was then centrifuged at 13000rpmfor 10 minutes. Compact white pellet was formed.

6. The spin column was placed in the collection tube and the supernatant was quickly passed through the spin column by giving a spin at 10000rpm for 1 minute. The flow through was discarded.
7. The spin column was then placed in the same collection tube and washed with 750µl of buffer W by giving a spin at 10000rpm for 1 minute. The flow through was discarded.
8. The spin column was placed in same collection tube and centrifuged for an additional 4 minutes.
9. DNA was eluted by dispensing 35µl of sigma water at the centre of the columns which is then kept for 1 minute and thereafter centrifuged at 12000rpm for a minute.

3.2.1.12 SDS-Polyacrylamide Gel Electrophoresis

The acrylamide percentage in SDS-PAGE depends upon the size of the target protein in the sample. The composition of the resolving and stacking gel is shown in table below.

Gel volume	**5ml**	**4ml**
Components	**Resolving gel (12%)**	**Stacking gel (5%)**
Water	1.6	2.7
Acrylamide/Bisacrylamide (30%/0.8% w/v)	2.0	0.67
1.5M Tris (pH 8.8)	1.3	-
1.0M Tris (pH 6.8)	-	0.5
10% SDS	50µl	40
10%APS	50µl	40
TEMED	3µl	4

Table 5: Composition of Resolving and stacking gel for SDS-PAGE

1. The two gel plates were clamped in the casting frames on the casting plates.
2. The resolving gel solution was prepared (as described above) and swirled gently but thoroughly.
3. Appropriate amount of the gel was pipetted into the gaps between the two plates.
4. Isopropanol was overlaid onto the gel surface to ensure flat interface between the resolving and stacking gel.
5. The gel was allowed to stand for 20-30 min for gelation.
6. The stacking gel was prepared. In the meanwhile, isopropanol was blotted out using a filter paper.

7. The stacking gel was then pipetted until overflow.
8. The well forming comb was inserted without trapping air beneath the teeth.
9. The gel was allowed to stand for 20-30 min for gelation.

Sample preparation

The protein samples were prepared by mixing them with SDS loading dye. The samples were heated in boiling water for 5 minutes followed by loading on the gel along with the protein ladder.

After completion of the electrophoresis run, the gel is kept in staining solution followed by destaining overnight.

Components	Volume
1M Tris-Cl (pH-6.8)	2.5
SDS	6.6g
Glycerol	5ml
B-mercaptoethanol	350µl
Bromophenol Blue	10mg

Table 6: Composition of 5X loading dye

Components	Volume
Tris-base	15.1g
Glycine	94g
SDS	5g
Volume made up to 1L with dH20	

Table 7: Composition of 5X running buffer

Components	Volume
Methanol	90ml
Glacial acetic acid	20ml
Water	90ml
Coomassie brilliant blue	0.5g
Filtered through Whattman No.1 filter paper	

Table 8: Composition of staining solution

Components	Volume
Methanol	90ml
Glacial acetic acid	20ml
Water	90ml

Table 9: Composition of destaining solution

3.2.1.13 Protein extraction and Solubilisation by Urea Treatment

1. Cell lysate was centrifuged at 16000rpm, 4^0C for 20 minutes.
2. The pellet was resuspended in 10ml wash buffer containing 1% Triton X-10 and 1M urea per gram cell weight and incubated at room temperature for 5 minutes.
3. The lysate was centrifuged at 16000rpm, 4^0C for 20 minutes.
   The inclusion body pellet was resuspended in 8M urea extraction buffer including 5mM PMSF and 100µl bacterial protease inhibitor, keeping it for stirring at room temperature for 1 h as described by Majee *et al.* 2004.
4. After solubilisation, the lysate was centrifuged at 8500rpm, 4^0C for 30 minutes
5. The supernatant was then diluted tenfold with extraction buffer and dialysed overnight in urea free buffer at 4^0C
6. The dialysate was then centrifuged at 8500rpm, 4^0C for 20 minutes

3.2.1.14 Purification by Ni-NTA affinity columns

1. The column was washed with 5ml distilled water using a 5 ml syringe and equilibrated with binding buffer.
2. The dialysed sample was then loaded on to the column and the effluent collected in 1.5ml micro-centrifuge tubes.
3. The column was then washed with 10ml of binding buffer to remove unbound protein
4. Bound protein was then eluted by passing elution buffer through the column and fractions collected in 1.5ml micro-centrifuge tubes
5. Active fractions were pooled together and dialysed against buffer A to remove remaining imidazole.

3.2.1.15 Protein Quantification

1. 10µl of the protein sample was mixed with 890µl of distilled water to which 100µl of Bradford reagent was added.
2. This was followed by incubation at room temperature for 5 minutes.
3. The absorbance was read at 595 nm. The concentrations of the samples were obtained from the standard curve.

3.2.1.16 Biochemical Characterisation of Recombinant CaIMPL1 and CaIMPL2

IMP ASSAY

Measurement of inorganic phosphate by Malachite green method

Malachite

6ml of conc.H2S04 was added in 30ml water. Solution was cooled to room temperature, 2.9 ml of malachite green added to the solution and stored in amber bottle at room temperature

Ammonium Molybdate

7.5% ammonium molybdate was added to warm water

Tween-20

Prepare a11 %(v/v) Tween-20 solution helps in stabilisation of colour of the solution.

Total No. of reactions	Total volume of malachite green	Malachite green in H_2S0_4	Ammonium Molybdate	11% Tween-20
6*200µl	1.27ml	1ml	0.25ml	20µl

Table 10: IMP Reaction mixture

Tris-Cl (pH-8) (0.1M)	50mM	50µl
DL-Inositol-1-Phosphate(1mM)	30µM	3µl
MgCl2(100mM)	3mM	3µl
Purified IMPL	10-20µg	10µl
Autoclaved MilliQ	To make up volume upto	100µl

IMP activity was assayed by colorimetric estimation of released inorganic phosphate after enzymatic hydrolysis of L-myo inositol- 1- phosphate with Malachite Green.

1. 2µg of purified Inositol-1-phosphatase was added in the reaction mixture and thereafter, incubated at 30^0C for 1hour.
2. After 1 hour, 700µl distilled water and 200µl Malachite Green solution was added and incubated for 10 min at room temperature for stabilisation of the developed colour
3. Absorbance was read at 630nm. The sample concentrations were obtained from the phosphate standard curve.

3.2.1.17 Agrobacterium mediated Plant Transformation by Floral Dip Method

1. Arabidopsis plants were grown in pots till onset of flowering followed by clipping of first bolts and already formed siliques.
2. Meanwhile, the plant constructs were transformed in *Agrobacterium* cells
3. The cells were then grown in 250 ml liquid broth supplemented with rifampicin and kanamycin for 16 - 24 hrs at 28°C.Centrifugation of cells was carried out at 6000 rpm for 5-7 min at room temperature.
4. Cells were resuspended in 600 ml of 5 % sucrose solution once A600 reaches 0.8. Silwet L-77 was added to a concentration of 0.03 % constantly mixing it well
5. Plant parts above the soil were immersed in *Agrobacterium* solution by inverting the pot for 2 min, followed by incubation in dark for 24 hours to maintain high humidity.
6. Transformants were selected using antibiotic or herbicide selectable marker.

3.2.1.18 Generation of plants over-expressing CaIMPL1 and CaIMPL2

3.2.1.18.1 Vector construction for over-expression construct

For plant over-expression, pCAMBIA 1301 modified vector was used carrying a 35S promoter. *CaIMPL1* and *CaIMPL2* were cloned under 35S promoter

3.2.1.18.2 Generation of Transgenic lines

Arabidopsis wild type plants would be used for *Agrobacterium* mediated transformation by floral dip method (section 3.2.1.20).

3.2.1.19 Gateway Cloning For Sub-cellular Localisation Study

Cloning in entry vector

For cloning into entry pENTD-TOPO vector, the pJET: *CaIMPL1* and pJET: *CaIMPL2* plasmids were amplified with a set of gene specific primers. The forward PCR primer contains the sequence, CACC, at the 5'end of the primer. The 4 nucleotides, CACC, base pair with the overhang sequence, GTGG present in the pENTD-TOPO vector. The PCR reaction was set up as follows:

Components	Volume(µl)
PCR Master Mix	22.5
Forward primer	1
Reverse primer	1
Template	0.5

Table 11: Reaction mixture for PCR Amplification

PCR parameters

Hot Start	54^0C	3 min	1 cycle
Denaturation	94^0C	30sec	35 cycles
Annealing	62^0C	25sec	
Extension	72^0C	2min 40sec	
Hold	4^0C	∞	Hold

The PCR product thus obtained was eluted (as described in section 3.2.1.5) and cloned in the entry vector pENTD-TOPO (Table 12) by a recombinant based method and transformed into *E.coli DH5α* cells. Cells were plated onto LB+Kanamycin plate, and incubated at 37^0C overnight.

3.2.1.20 Bioinformatics analysis of *CaIMPL1* and *CaIMPL2*

The amino acid sequence of CaIMPL1 and CaIMPL2 for different plant species were retrieved from the JGI database. The amino acid sequence comparison and alignments were performed with local versions of BLAST and MEGA software

Chapter 4

Isolation and Cloning of *CaIMPL1* and *CaIMPL2* cDNA from Chickpea

INTRODUCTION

The *IMP* gene was first cloned from bovine brain tissue and there have been excellent studies on the IMP enzyme from several plant and animal tissues (Chen and Charalampous, 1966: Eisenberg, 1967; Hallcher and Sherman, 1980; Majerus et al., 1999). The single documented biosynthetic route of myo-inositol begins with the conversion of glucose-6-phosphate to D-myo-inositol-3-phosphate catalysed by myo inositol phosphate synthase and ends with de-phosphorylation by myo-inositol monophosphatase (IMP) to generate free inositol (Loewus and Murthy,2000).

Arabidopsis *imp* mutants have only a 30% reduction in inositol content, suggesting that other redundant enzymes function during *myo*-inositol synthesis. IMPL proteins are good candidates for such enzymes, as they are the two most closely related proteins in the Arabidopsis genome.IMPL1 and IMPL2 contain the conserved inositol P domain found in all characterised IMPs. Indeed, all plants queried contain multiple *IMP-like (IMPL)* genes, closer in amino acid sequence identity to prokaryotic IMPs (Torabinejad and Gillaspy, 2006).

The molecular regulation and characterisation of the gene encoding IMPL is limited to Arabidopsis only. No IMPL gene has been cloned and characterised from legumes so far. Thus, in the first part of the study, two genes/cDNA encoding IMPL *(CaIMPL1 and CaIMPL2)* were isolated and cloned from chickpea. Genome structures were analysed and exon-intron boundaries were determined. The deduced amino acid sequences were analysed for their phylogenetic relationship with IMPL from other organism and plant sources.

METHODOLOGY

Isolation of total RNA

Total RNA was isolated from chickpea seedling using TRI reagent according to the protocol mentioned in section 3.2.1.1. The isolated RNA was then purified of protein and DNA contamination and cDNA subsequently synthesised.

cDNA synthesis

The corresponding cDNA was synthesised using oligodT anchor primer and reverse transcriptase according to the protocol mentioned in section 3.2.1.3.

PCR amplification of *CaIMPL1* and *CaIMPL2* using gene specific primers

Gene specific primers were designed for amplification of full length cDNA of *CaIMPL1* and *CaIMPL2* from the total cDNA. These primers were designed based on the *CaIMPL1*, *CaIMPL2* sequence information obtained by homology searching against Arabidopsis *IMPL1, IMPL2*. The sequences of *AtIMPL1* and *AtIMPL2* were retrieved from TAIR database and subsequently BLAST against the target organism *Cicer arietinum* in the Legume Information System database.

Cloning into pJET vector

The *IMPL1, IMPL2* amplicons thus obtained were run on 0.8% agarose gel. The bands were then eluted from the gel using gel extraction kit according to the manufacturer's protocol (section 3.2.1.5). The eluted products were ligated into pJET 1.2 (blunt end) cloning vector. The ligation mixture was incubated for 30 minutes at room temperature and directly transformed into *E.coli DH5α* competent cells (Section 3.2.1.9). Cells were plated onto LB agar ampicillin plates and incubated overnight at 37^0C. Transformed colonies were streaked on LB ampicillin plates and screened for positive clones by colony PCR (as described in section 3.2.1.10) and restriction digestion (as described in section 3.2.1.6). Plasmid DNA was isolated (as described in section 3.2.1.11) from the positive colonies, further confirmed by sequencing.

Restriction digestion of pJET: *CaIMPL1* and pJET: *CaIMPL2*

The pJET: *CaIMPL1* and pJET: *CaIMPL2* were digested using restriction enzymes, the restriction sites of which were added within the primers used for amplification. The pJET: *CaIMPL1* was digested with NdeI/XhoI restriction enzyme pair and pJET: *CaIMPL2* by NcoI/XhoI enzyme pair. The reaction mix was incubated at 37^0C for 1 hour. Thereafter, the samples were analysed on 0.8% agarose gel (as described in section 3.2.1.4).

Chapter 5

Bacterial Over-expression and Purification of CaIMPL1 and CaIMPL2

INTRODUCTION

Among plants, only Arabidopsis IMPL proteins have been characterised till date. Thus, in the second part, we were interested to study the biochemical properties of the IMPL enzymes in chickpea.

Initially, the proteins were expressed with a C-terminal hexa- histidyl tag in E.coli. The bacterially over-expressed recombinant proteins were purified using affinity chromatography. These purified proteins were then analysed for their enzymatic activity.

METHODOLOGY

Sub-cloning of *CaIMPL1* and *CaIMPL2* cDNA into pET28a vector

pJET: *CaIMPL1* and pJET: *CaIMPL2* plasmids were digested using restriction enzymes NdeI/XhoI and NcoI/XhoI respectively, followed by heat inactivation at 65^0C for 20 minutes. The digested products were run on 0.8% agarose gel (as described in section 3.2.1.4). The required bands were eluted from the gel using gel extraction kit according to the manufacturer's protocol (as described in section 3.2.1.5). Similarly, pET28a empty vector was also digested by NdeI/XhoI and NcoI/XhoI restriction enzymes for incorporation of the two inserts respectively, followed by heat inactivation at 65^0C for 20 minutes and alkaline phosphatase treatment. The digested empty vector was then purified. The ligation mixture was incubated at 22^0C for one hour and directly transformed into *E.coli DH5α* competent cells (as described in section 3.2.1.9). Cells were plated onto LB+ Kanamycin plates, and incubated at 37^0C. The transformed colonies were streaked on LB kanamycin plates and screened for positive clones by colony PCR (as described in section 3.2.1.10). Then plasmid DNA was isolated (as described in section 3.2.1.11) from the positive clones which were again confirmed by restriction digestion (as described in section 3.2.1.6). These plasmids were further confirmed by sequencing.

Bacterial over-expression of CaIMPL1 and CaIMPL2

For bacterial over-expression of CaIMPL1 and CaIMPL2, plasmids isolated from the confirmed clones in DH5α cells were transformed into BL21 (DE3) and Lemo21 *E.coli* competent cells. Cells were plated onto LB+Kanamycin plates, and incubated at 37^0C. The transformed colonies were streaked on LB kanamycin plates and screened for positive clones by colony PCR. 100µl of the primary culture of the positive clones of CaIMPL1 and CaIMPL2 was inoculated in 5ml LB+Kan culture media, incubated at 37^0C for 1hr 30 minutes. The culture was then induced by 0.5mM IPTG as soon as the absorbance reached 0.6-0.8 at 600nm.

Cells were harvested using centrifugation, sonicated at 0.6 cycles and 70% amplitude and finally analysed on 12% SDS-PAGE (as described in section 3.2.1.12).

Solubilisation of the recombinant protein

1ml of primary culture of the positive clones of *CaIMPL1* and *CaIMPL2* were inoculated in 200ml LB+Kanamycin culture and incubated at 37^0C, 180rpm for 2hrs 30 minutes. The cultures were then induced with 0.5mM IPTG when the absorbance at 600nm read 0.6. These cultures were then incubated at 18^0C overnight. The cells were harvested by centrifugation, followed by sonication and extraction (as described in section 3.2.1.13). Solubilisation of protein samples was carried out by Urea solubilisation (as described in section 3.2.1.13)

Purification of the dialysed expressed protein sample using nickel charged affinity columns

The dialysate was then centrifuged at 8500rpm, 4^0C for 20 minutes. The supernatant of the two recombinant proteins were then filter sterilised using 0.45micron filter. 2mM of imidazole was added to each of the protein samples and purified using nickel charged affinity columns (as described in section 3.2.1.14).. The purified fractions were checked on 12% SDS-PAGE.

Biochemical Characterisation of CaIMPL1 and CaIMPL2

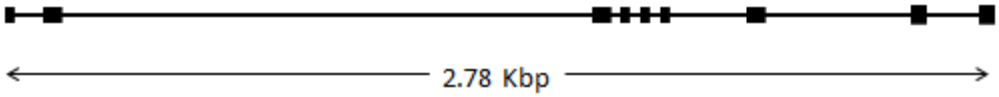
CaIMPL1 and CaIMPL2 purified proteins were quantified using Bio-Rad protein estimation kit and activity measured using Malachite Green method (as described in section 3.2.1.16).

RESULT AND DISCUSSION

6.1 Isolation and Cloning of *CaIMPL1*and *CaIMPL2* cDNA

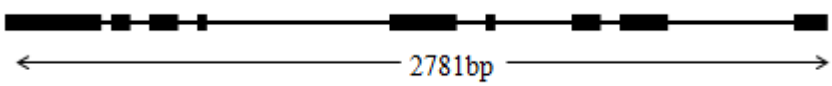
6.1.1 Identification of IMPL gene in Chickpea genome

In order to identify Inositol monophosphatase like *(IMPL)* genes in chickpea genome, coding sequences of model plant Arabidopsis *IMPL* genes were retrieved from TAIR database (https://www.arabidopsis.org/). Total of two Arabidopsis IMPL genes were found. *AtIMPL1 and AtIMPL2* sequences with accession number *At1g31190* and *At4g39120* were blast against the target organism *Cicer arietinum* using the blastn software in the Legume Information System database. Similarly, CaIMPL1 and CaIMPL2 protein sequences were retrieved using the blastp software in the LIS database. The full length sequence of *CaIMPL1* including the 5' and 3'UTR was found to be 1329 bp in length, containing an open reading frame of 1044 bp encoding a polypeptide of 347 aa.



Diagrammatic representation of CaIMPL1 genomic sequence

The full length composite sequence of CaIMPL2 including the 5' and 3'UTR is 1234 bp in length, containing an open reading frame of 894 bp encoding a polypeptide of 297 aa.



Diagrammatic representation of CaIMPL2 genomic sequence

Blast analysis with the recently released chickpea genome sequence (Varshney et al., 2013) revealed that *CaIMPL1* gene is located on chromosome 2 whereas *CaIMPL2* gene is located on chromosome 3 of *Cicer arietinum*.

6.1.2 Phylogenetic analyses of CaIMPL1 and CaIMPL2

ClustalW was used for the alignment of the multiple amino acid sequences.

```
AtIMPL2         MRLHNRDLSIQTFSVVDDWLFRELPTHSKMLAQSHFFSKSFDLIPPQSPALRSANPSLRI
CaIMPL2         ------------------------------MLSQCHLHCYSNNLS------IRSPKLRLRA
AtIMPL1         ------------------MGRSLIFSGNMSLRISHLPRSSLPLQNPISGRTVNRTFRYRC
CaIMPL1         --------------------MSIVFSAASNLSWHKDCRQSSPPIGSWRLKSRIQSCKN--
                                     *   :    *             .


AtIMPL2         SSSYSNSRLSFLSSSAIAVPVSRRRFCLTMASNSKRPNISNESPSELSDTELDRFAAVGN
CaIMPL2         MSSSS-------------------------------------------SPHQFNHFADVAN
AtIMPL1         TRILSNSFKS----------TTRLQTKAVLSEVSDQTRYPRIGAKTTGTISPAHLLEVVE
CaIMPL1         ------------------------------SLQSDIYTQHRVGARSTGPIQPTHLIQVAT
                                                          .  .  ::  *

                                                                 Motif C

AtIMPL2         ALADASGEVIRKYFRKKFDIVDKDDMSPVTIADQMAEEAMVSIIFQNLPSHAIYGEEKGW
CaIMPL2         KAADAAGDVIRKYFRKNFDIIHKHDLSPVTIADQTAEEAMVSIILDNFPSHAVYGEEKGW
AtIMPL1         LAAKTGAEVVMEAVNKPRNITYKGLSDLVTDTDKASEAAILEVVKKNFSDHLILGEE-GG
CaIMPL1         TAAQTGAQVVMDAVNKPRNITYKGLTDLVTETDKMSEAAILEVVKKNFDDHLILGEE-GG
                  *.:..:*: . ..*  :*  *   . ** :*: :* *::.:: .*: .* : *** *

                           Motif A

AtIMPL2         RCKEESADYVWVLD-PDGTKSFITGKPVFGTLIALLYKGKPILGLIDQPILK-ERWIG--
CaIMPL2         RCRQDSADYVWVLDPIDGTKSFITGKPLFGTLIALLQNGTPILGIIDQPVLR-ERWIG--
AtIMPL1         IIGDSSSDYLWCIDPLDGTTNFAHGYPSFAVSVGVLYRGNPAAASVVEFVGGPMCWNTRT
CaIMPL1         IIGEAASDYLWCIDPLDGTTNFAHGYPSFAVSVGVLYRGNPAAATVVEFVGGPMCWNTRI
                   : ::**:* :*  ***..*  * * *.. :.:* .*.*  . : : :     *


AtIMPL2         ---MNGRRTKLNGEDISTRSCPKLSQAYLYTTS----PHLFSEEAEKAYSRVRDKVKVPL
CaIMPL2         ---MTGKRTTLNGQEVSTRTCADLSQAYLYTTS----PHLFSGDAEEAFIRVRDKVKIPL
AtIMPL1         FSATAGGGALCNGQKIHVSKTDAVERALLITGFGYEHDDAWSTNMELFKEFTDVSRGVRR
CaIMPL1         FTATAGGGAFCNGQRIHVSATNQVEQSLLVTGFGYEHDEAWATNIELFKEFTDVSRGVRR
                     *  :  **: : .     :.:: * *       . :: : *     .  .  :

                                        Motif B
AtIMPL2         YGCDCYAYALLASGFVDLVIESGLKPYDFLALVPVIEGAGGTITDWTGKRFLWEASSSAV
CaIMPL2         YGCDCYAYALLSSGFVDLVVESGLKPYDFLALVPVIEGSGGVITDWEGHQLRWEASPLSI
AtIMPL1         LGAAAVDMCHVALGIAESYWEYRLKPWDMAAGVLIVEEAGGAVTRMDGGKFSVFDRSVLV
CaIMPL1         LGAAAVDMCHVALGIVEAYWEYRLKPWDMAAGVLMVEEAGGTVSRMDGGKFCVFDRSVLV
                 *. .   . :: *:.:   *  ***:*: * * ::* :**.::   * ::     .  :


AtIMPL2         ATSFN------VVAAGDSDIHQQALESLEWH---------
CaIMPL2         AISFN------VVAAGDKQIHQQALDSLQR----------
AtIMPL1         SNGVLHPKLLERIAPATENLKSKGIDFSLWFKPEDYHTEL
CaIMPL1         SNGVLHTELLERIGPATEELKSKGIDFSLWYKPEDYRADV
                : ..        :... .::::.:.::
```

Figure B(i): Multiple sequence analysis of IMPL sequences from Arabidopsis and Chickpea using the Clustal-W tool. Three characteristic signature motifs (Motif 'A' - DPIDGT, Motif 'B' - WDXAAG, and Motif 'C' - GEES) of the phosphatase super family is highlighted

Figure B (ii): Radial format phylogenetic tree describing the evolutionary relationship between prokaryotic IMPL, yeast IMPL, Human IMPL proteins and Fabaceae family IMPL1 and IMPL2 proteins.

The amino acid sequence of prokaryotic and eukaryotic (human) IMPL protein was retrieved from NCBI protein database. The amino acid sequence of different plant IMPL proteins were obtained from the JGI database. The phylogenetic tree was constructed based on the neighbour joining method using the MEGA (Molecular Evolutionary Genetics Analysis) software.

From the phylogenetic tree generated, it can be observed that the IMPL1 protein of the Fabaceae family is most closely related to member of the Brassicaceae family, *Arabidopsis* IMPL1 protein followed by Human IMPA1 and IMPA2 protein, yeast IMP1and IMP2 protein and the prokaryotic IMPL, *Halobacterium* SuhB protein. As compared to the IMPL1 protein of the Fabaceae family, the IMPL2 protein of the Fabaceae family is distantly related to the Human IMPL protein and yeast IMP proteins, although closely related to *Arabidopsis* IMPL2 protein and prokaryotic IMPL, *E.coli* CysQ protein. However, *E.coli* SuhB protein is approximately equidistant to both IMPL1 and IMPL2 protein of the Fabaceae family. Close relatedness suggests that not much of genetic change has occurred over time.

6.1.3 Primer designing

Gene specific primers for IMPL genes were designed based on the nucleotide sequencing of CaIMPL obtained from the Legume Information System database.

	Primers
CaIMPL1(F)	CATATGATGTCAATTGTATTCTCCGCAGC
CaIMPL1(R)	CTCGAGAACGTCTGCCCTGTAGTCCTC
CaIMPL2(F)	CCATGGATGTCAATTGTATTCTCCGCAGC
CaIMPL2(R)	CTCGAGATGGTGTCACAGTGCCATCT

Table 12: List of gene specific primers used for amplification of full length CaIMPL1 and CaIMPL2. The underlined portion indicate the restriction sites for NdeI/XhoI and NcoI/XhoI

6.1.4 Isolation of IMPL genes

To isolate full length CDS of *IMPL1* and *IMPL2* from chickpea, Total RNA was isolated from chickpea seedling according to the protocol mentioned in section 3.2.1.1 and the corresponding cDNA was synthesised using oligo dT anchor primer and reverse transcriptase according to the protocol mentioned in section 3.2.1.3.

6.1.5 Cloning into pJET vector

Gene specific primers were designed for amplification of full length cDNA of *CaIMPL1* and *CaIMPL2* from the total cDNA. The *IMPL1, IMPL2* amplicons thus obtained were checked on 0.8% agarose gel. The bands were then eluted from the gel using gel extraction kit (as described in section 3.2.1.5). The eluted products were ligated into pJET 1.2 (blunt end) cloning vector.

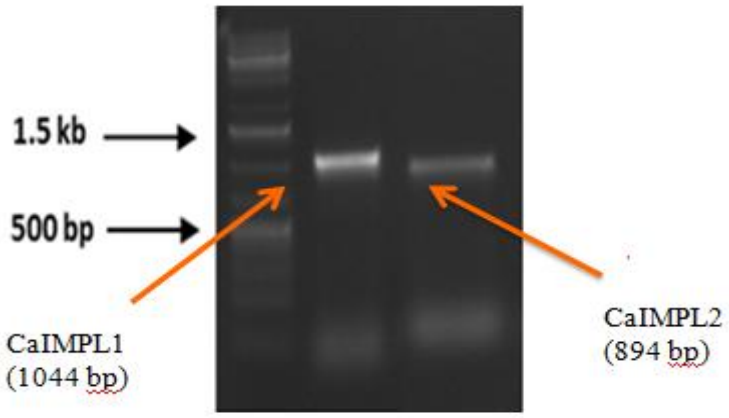


Figure B1: 0.8% Agarose gel showing IMPL1 and IMPL2 amplicons

The ligation mixture was kept at room temperature for 30 minutes and directly transformed into *E.coli DH5α* competent cells (as explained in section 3.2.1.9). Cells were plated onto LB agar ampicillin plates and incubated overnight at 37^0C. Transformed colonies were streaked on LB ampicillin plates. Colony PCR was used for screening of positive clones (as described in section 3.2.1.10) and restriction digestion (3.2.1.6). Plasmid was isolated (as described in section 3.2.1.11) from the positive colonies, further confirmed by sequencing.

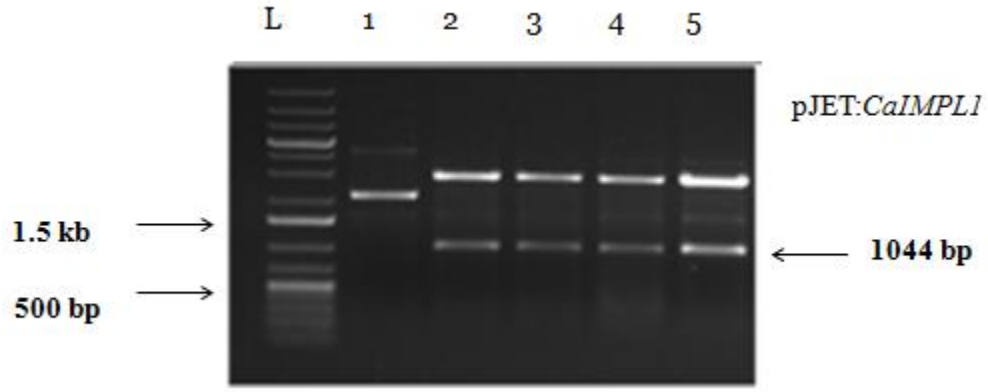


Figure B2:0.8% Agarose gel showing restriction digestion of pJET:*CaIMPL1* by NdeI/XhoI

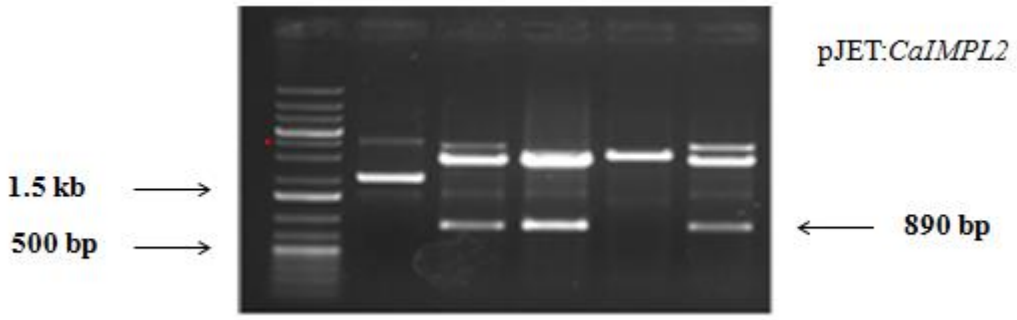


Figure B3:0.8% Agarose gel showing restriction digestion of pJET:*CaIMPL1* by NcoI/XhoI

The nucleotide sequence obtained from the sequencing of pJET: *CaIMPL1* and pJET: *CaIMPL2* were compared with the sequences obtained from the Legume Information System database using the DNA Star SeqMan II software. Based on the results of sequence analysis, positive clones were identified.

6.2 Bacterial Over-expression and purification of CaIMPL1 and CaIMPL2

Heterologous protein expression requires the insertion of the desired DNA fragment (open reading frame) into an expression vector and transformation of the resulting plasmid into host bacterial cell. The cells are then induced to obtain proteins ready for purification and characterisation.

6.2.1 Sub-cloning of CaIMPL1 and CaIMPL2 into pET28a expression vector

For sub-cloning the positively selected clones of *CaIMPL1* and *CaIMPL2* were digested with NdeI/XhoI and NcoI/XhoI restriction enzymes respectively. The bands corresponding to the correct insert size were eluted from the gel (as described in section 3.2.1.5) and ligated to the pET28a vector (digested with the same enzyme pair). The ligation mixture was incubated for 30 min and then directly transformed into the host cell DH5α. Cells were plated onto media plates supplemented with kanamycin, and incubated at 37^0C. The transformed colonies were streaked on LB kanamycin plates. Colony PCR was used for screening of positive clones (as described in section 3.2.1.10). Then plasmid DNA was isolated from (as described in section 3.2.1.11) from the positive clones which were again confirmed by restriction digestion (as described in section 3.2.1.6). These plasmids were further confirmed by sequencing.

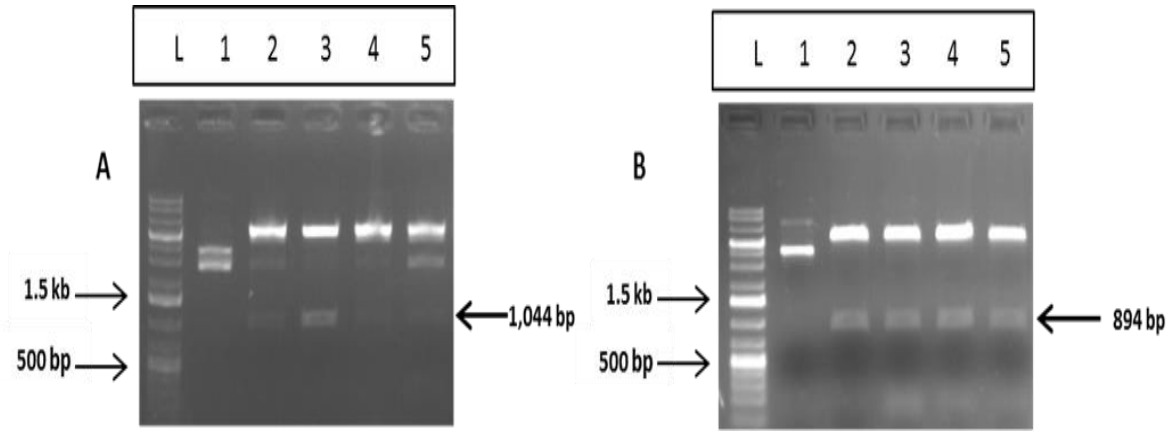
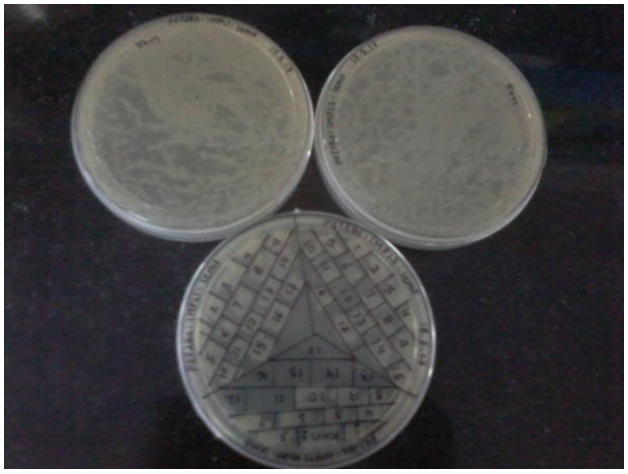


Figure B4: 0.8% Agarose gel showing Restriction digestion of (A) pET28a:*CaIMPL1* and (B) pET28a: *CaIMPL2*

6.2.2 Transformation of the resulting plasmid into bacterial host strain *E.coli* Lemo21 *and* BL21(DE3)

The confirmed plasmids were subsequently transformed into bacterial host strain *E.coli* Lemo21 and BL21 (DE3) cells for protein expression.

A



B

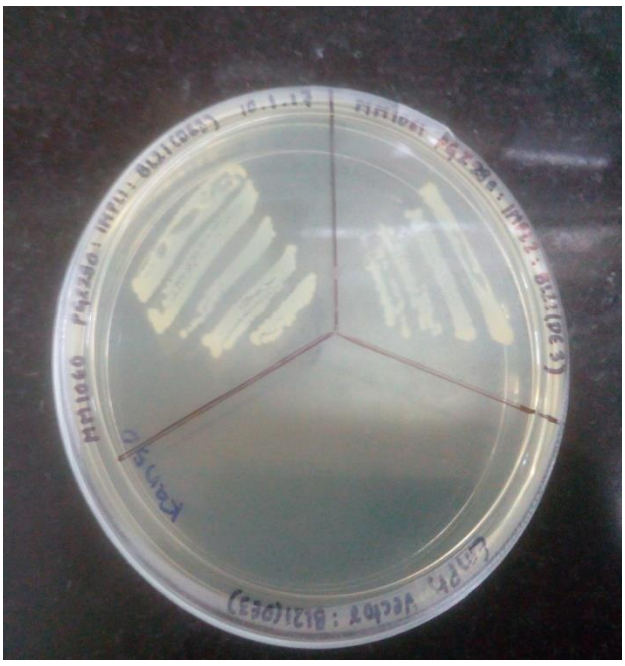


Figure B5: (A) Transformed colonies of pET28a:*IMPL1*: Lemo and pET28a:*IMPL2*: Lemo and Patch plate of the colonies (B) Patch plate of pET28a:*IMPL1*: BL21 (DE3) and pET28a:*IMPL2*: BL21 (DE3)

6.2.3 Bacterial Expression of CaIMPL1 and CaIMPL2

For bacterial expression of the proteins, the selected transformed colonies of each construct were grown in LB broth supplemented with suitable antibiotic. The expression of the recombinant proteins was induced by adding IPTG. Induced cultures were further grown for 6 hour at 25^0C before harvesting the cells. The cells were sonicated and particulate and soluble fractions collected separately. These fractions were then analysed on 12% SDS-PAGE

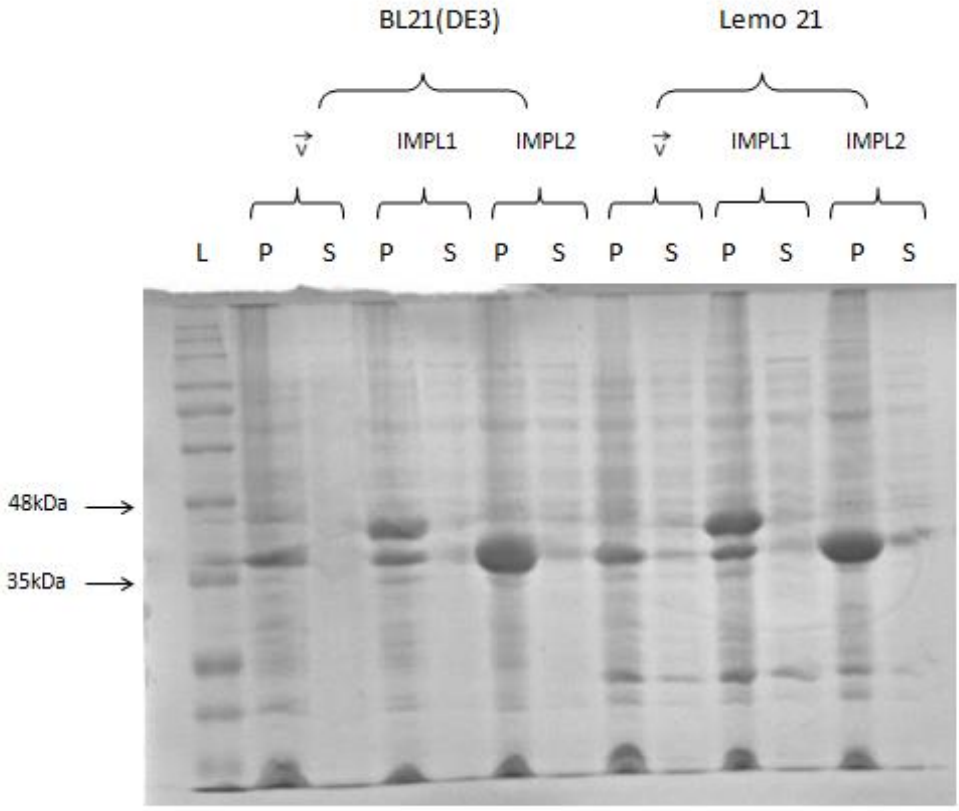
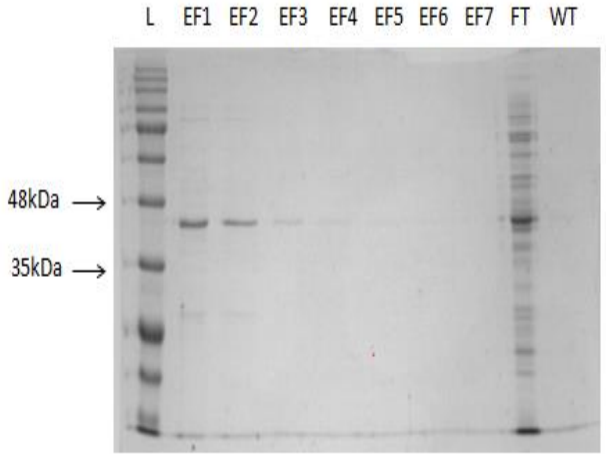


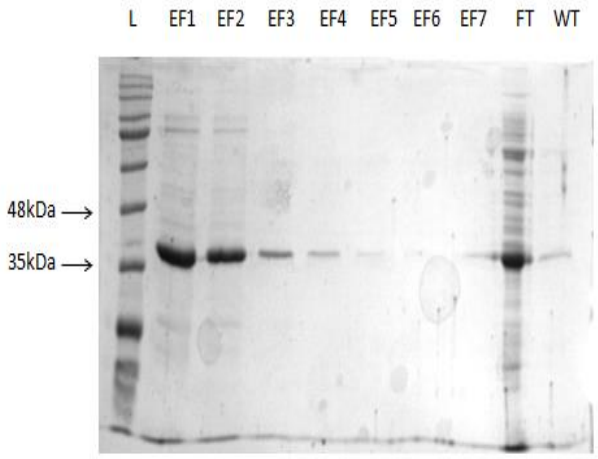
Figure B6: Analysis of bacterial over-expression of CaIMPL1 and CaIMPL2 [L-Ladder, P-Pellet, S-Supernatant, V-pET28a empty vector]

Both, recombinant IMPL1 and IMPL2 were found to migrate slightly slower than expected given their predicted molecular masses of 38.9 and 33.4kDa respectively (Figure B6). CaIMPL1and CaIMPL2 both expressed predominantly in the particulate fraction. In order to express the proteins in soluble fraction, induced cultures were incubated at low temperature of about 18^0C, overnight. However, both the proteins were expressed in the particulate fractions. This could possibly be due to the high level of expression of the recombinant proteins in *E.coli* resulting in accumulation as insoluble agglomerates *in vivo* as inclusion body pellet.

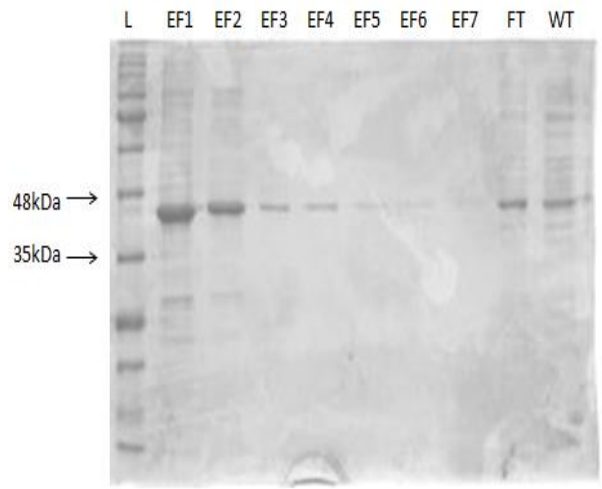
The particulate fractions were first solubilised in 8M urea and then dialysed before purification through affinity chromatography. Immobilised Metal Affinity Chromatography (IMAC) was performed to purify the bacterially expressed proteins using nickel charged affinity columns. Purified protein fractions were checked on 12% SDS-PAGE.



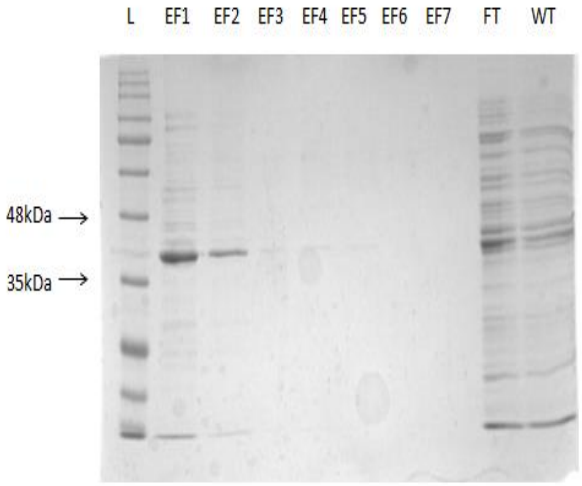
Purified fractions of pET28a:IMPL1:Lemo



Purified fractions of pET28a:IMPL2:Lemo



Purified fractions of pET28a:IMPL1:BL21(DE3)



Purified fractions of pET28a:IMPL2:BL21 (DE3)

Figure B7: Purified fractions of CaIMPL1 and CaIMPL2

Purified fractions were then pooled and dialysed to remove the imidazole. Thereafter the pooled fractions were used to study the biochemical properties of the proteins.

6.2.4 Quantification of protein using Bio-Rad protein estimation kit

The concentration of protein was determined with reference to a standard 1X BSA Solution. A series of dilutions of known protein concentration were prepared and the concentration of the unknown sample was then determined based on the standard curve. The standard curve was prepared by plotting absorbance (595nm) against the protein concentration (µg/µl). The standard curve was considered to be calibrated successfully at R^2=0.98

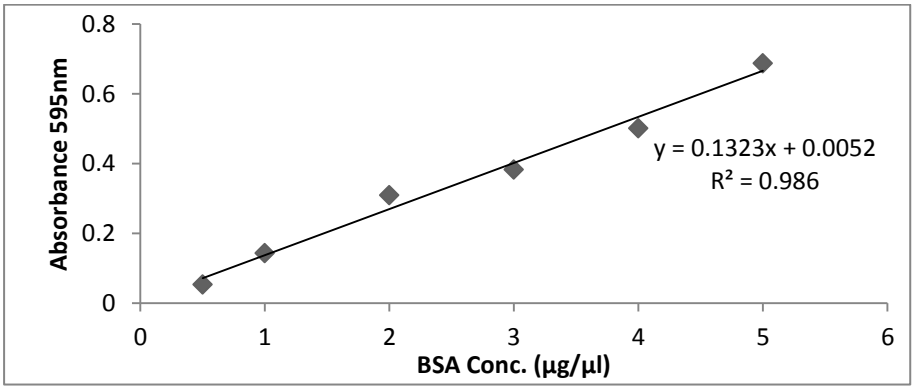


Figure B8: BSA standard Curve

The concentration of protein samples over-expressed and purified from two different host cells was found to be as follows:

Host Cell: Lemo21

Sample	OD595	Mean OD	Conc.(µg/10µl)	Conc.(µg/µl)
IMPL1	0.215	0.2115	1.602	***0.1602***
IMPL1	0.208			
IMPL2	0.067	0.11	0.833	***0.0833***
IMPL2	0.144			

Host Cell: BL21 (DE3)

Sample	OD595	Mean OD	Conc.(µg/10µl)	Conc.(µg/µl)
IMPL1	0.027	0.057	0.39	***0.039***
IMPL1	0.087			
IMPL2	0.126	0.1255	0.912	***0.0912***
IMPL2	0.125			

6.2.5 Biochemical characterisation of CaIMPL1 and CaIMPL2

Purified fractions were pooled to study the biochemical properties of the proteins.The enzymatic activity of CaIMPL1 and CaIMPL2 purified proteins was determined using Malachite Green method (as described in section 3.2.1.16).

6.2.5.1 Preparation of phosphate standard curve

The concentration of released inorganic phosphate by enzymatic reaction was determined with reference to a standard phosphate (KH_2PO_4) Solution. A series of dilutions of known phosphate concentration were prepared and the concentration of the unknown sample was then determined based on the standard curve. The standard curve was prepared by plotting absorbance (630 nm) against the protein concentration (µg/µl). The standard curve was considered to be calibrated successfully at R^2=0.985

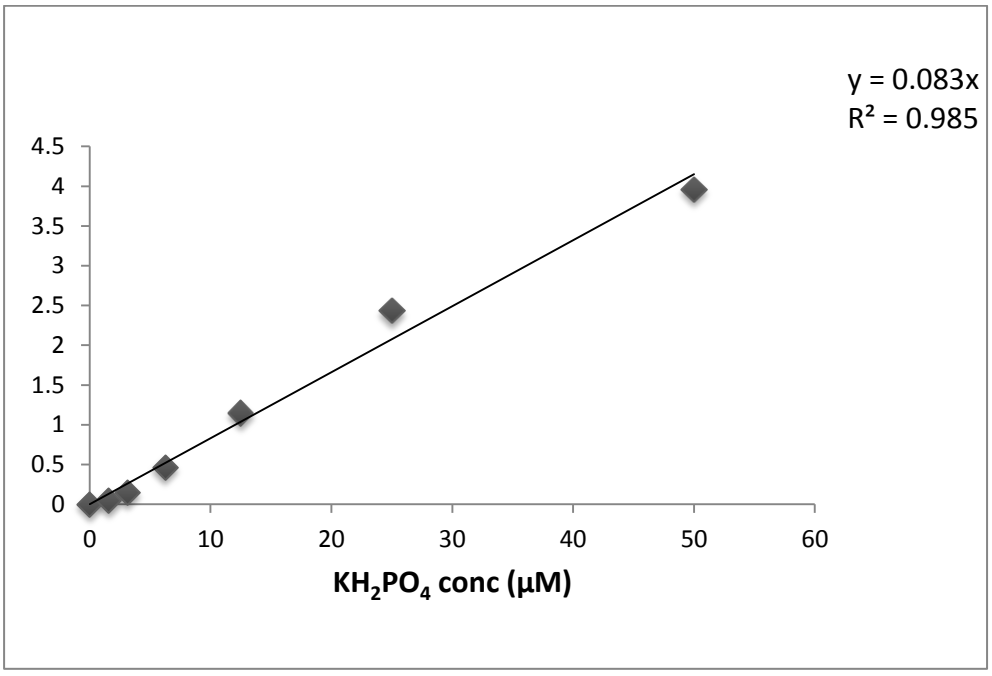


Figure B9: Phosphate Standard Curve

6.2.5.2 IMPL Assay

IMPL assay of the purified fractions of CaIMPL1 and CaIMPL2 was performed using the Malachite green method. The results are tabulated as follows

Incubation time: 60 min

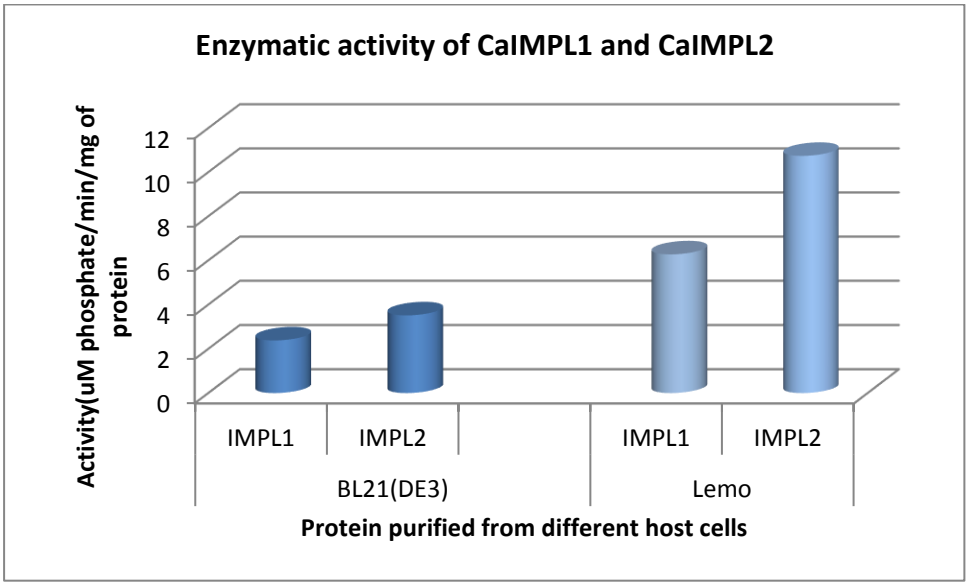
Concentration of protein: 2µg

Host Cell: Lemo21

Sample	**OD_{630}**	**Mean OD**	**Conc.(µM)**	**Activity(µM phosphate/min/mg of protein)**
IMPL1	0.049	0.0625	0.753	***6.275***
IMPL2	0.076			
IMPL1	0.12	0.107	1.28	***10.745***
IMPL2	0.094			

Host Cell: BL21 (DE3)

Sample	**OD_{630}**	**Mean OD**	**Conc.(µM)**	**Activity(µM phosphate/min/mg of protein)**
IMPL1	0.104	0.119	1.433	***2.389***
IMPL2	0.134			
IMPL1	0.140	0.175	2.108	***3.514***
IMPL2	0.210			



The comparative analysis of the enzymatic activity of CaIMPL1 and CaIMPL2 protein shows that proteins over-expressed and purified from Lemo*21* have higher activity as compared to those over-expressed and purified from BL21 (DE3). This could possibly be due to the slight variations in the features of the two host cells. Lemo21 offers tunable expression of clones. This results in more properly folded proteins which in turn contribute to higher bioactivity of protein.

6.4 Plant Transformation

For the expression of *CaIMPL1* and *CaIMPL2* in *Arabidopsis thaliana*, the required constructs were prepared using pCAMBIA1301 modified vector carrying 35S promoter and rbcs terminator. Initially the *CaIMPL1* and *CaIMPL2* cDNA was cloned in pJET vector and after selection of positive clones, it was sub-cloned into pCAMBIA1301 vector under 35S promoter at SacI/XbaI site. The resulting plasmids were transformed in *E.coli* DH5α cells. Cells were then plated on LB+Kan media and incubated at 37^0C overnight. The transformed colonies were then confirmed by restriction digestion followed by sequencing. Plasmids were isolated (as described in section 3.2. 1.11) from the positively selected clones and thereafter transformed in *Agrobacterium* GV3101. Cells were plated on LB+Kan+Rif media and incubated at 28^0C for 2 days. The transformed colonies (as shown in Fig. B10) were confirmed by colony PCR.

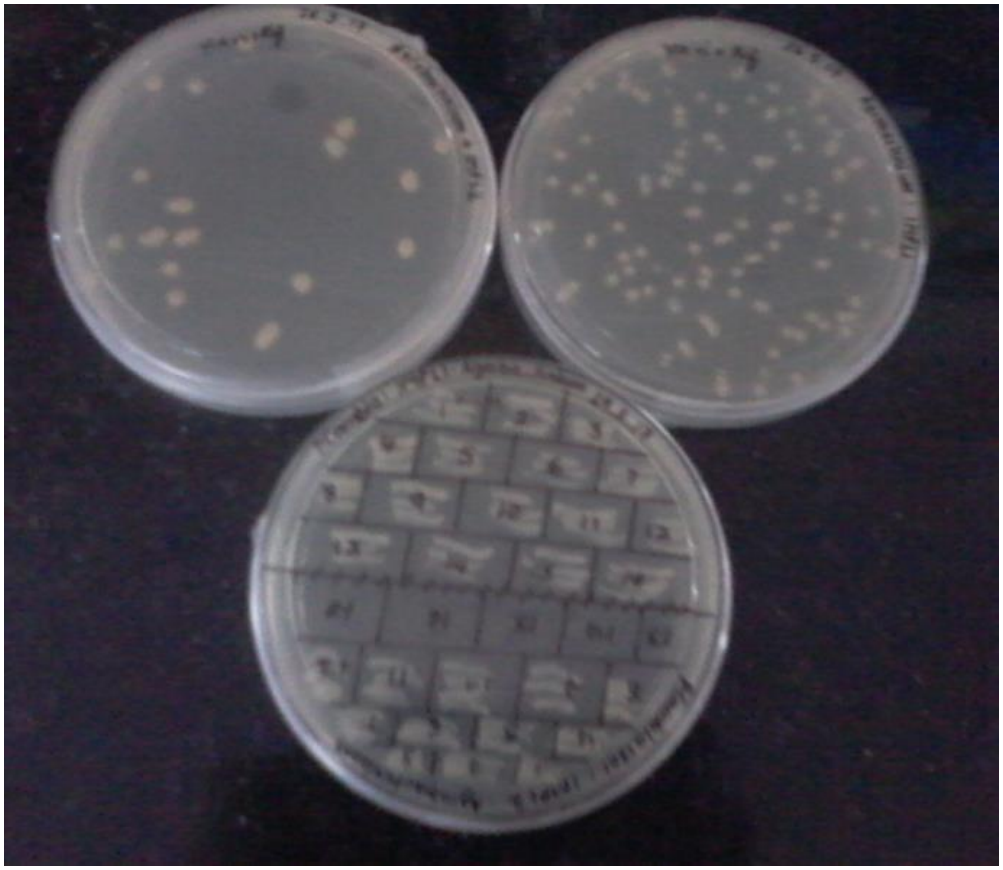


Figure B10: Transformed colonies of Pcambia1301:*IMPL1*: Agrobacterium and pCAMBIA1301:*IMPL2*: Agrobacterium.

6.4.2 Gateway cloning for sub-cellular localisation study

For localisation study of CaIMPL1 and CaIMPL2, the required constructs were prepared using Gateway cloning technique.

6.4.1 Cloning in entry vector

For cloning in entry vector, the pJET: CaIMPL1 and pJET: CaIMPL2 plasmids were amplified with a set of gene specific primers. The PCR product was then eluted and cloned in entry vector pENTD-TOPO vector by a recombination based method, subsequently transformed in *E.coli* DH5α cells (as shown in Figure B11) Plasmid was isolated from individual colonies and positive clones were screened by restriction digestion using restriction enzymes AscI/NotI and subsequently sequenced using gene specific primers.

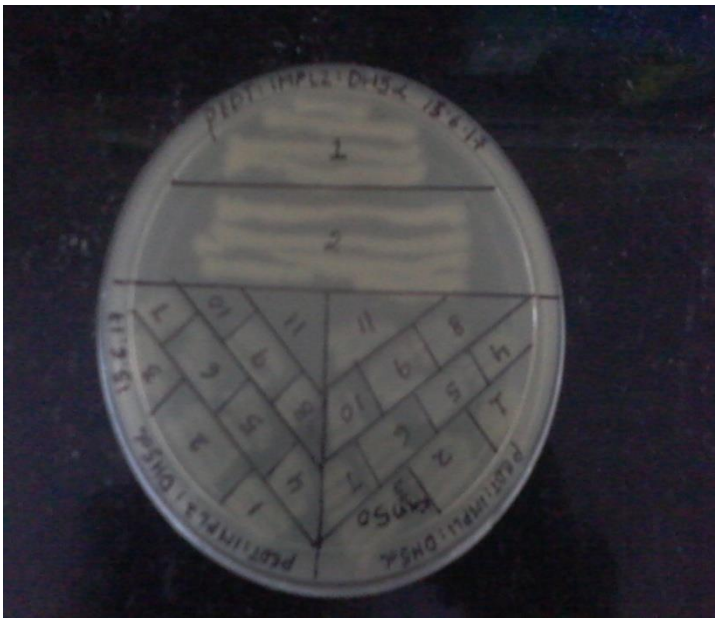


Figure B11: Patch plate of pEDT: *CaIMPL1* and pEDT: *CaIMPL2* transformed colonies

The sequencing result of pENTD-TOPO: *CaIMPL1* and *CaIMPL2* were analysed. Positive clones will be used for further studies.

CONCLUSION

Inositol plays an important role in the physiological well being of organisms. It is a versatile compound that generates diversified derivatives upon phosphorylation. The significance of these compounds lies in their dual functions as signals and as key metabolites under stress. In order to maintain proper function, the plant cell conserves a free pool of inositol through the *de novo* synthesis pathway and the recycling of phosphoinositide signalling molecules (Gillaspy, 2011). Several genes have been found to be crucial for synthesis and metabolism of inositol.

In this project, our objective was to study the enzymatic properties of bacterially over-expressed and purified CaIMPL1 and CaIMPL2. For that we isolated the total RNA from chickpea seedlings, synthesised the corresponding cDNA followed by amplification of full length cDNA for *CaIMPL1* and *CaIMPL2* using gene specific primers. The amplicons were cloned into the suitable vector and the positive clones confirmed by restriction digestion and sequencing. The positive clones of *CaIMPL1* and *CaIMPL2* were sub-cloned into expression vector for bacterial expression. Bacterial expression was carried out in two different host cells. Later, the proteins were purified using Ni-NTA affinity columns. The activity of the purified fractions was measured using the Malachite Green method.

We found that both the recombinant CaIMPL1 and CaIMPL2 were bioactive. The difference in enzymatic activity of CaIMPL1 and CaIMPL2 could possibly be attributed to the amino acid substitutions in the sequence motif. Also, CaIMPL1 and CaIMPL2 over-expressed and purified from Lemo21 had higher activity as compared to those over -expressed and purified from BL21 (DE3).

Further, for functional analysis, plant constructs have been prepared both for generation of transgenic lines and for sub-cellular localisation study.

REFERENCES

1. A.L. Majumder, M.D. Johnson, S.A. Henry, **(1997)** 1L-myo-Inositol-1-phosphate synthase, Biochim. Biophys. Acta **1348**, 245–256.

2. Atack JR, Broughton HB, Pollack S. **(1995).** Structure and mechanism of inositol monophosphatase. FEBS Letters **361**, 1–7

3. C.A. Brearley, D.E. Hanke, **(1996)** Metabolic evidence for the order of addition of individual phosphate esters to the *myo*–inositol moiety of inositol hexaphosphate in the duckweed *Spirodela polyrhiza*, Biochem. J. **314** 227–233.

4. Carman, G. M. & Kersting, M. C. **(2004)** Phospholipid synthesis in yeast: regulation by phosphorylation. *Biochem. Cell Biol.* **82**, 62–70**.**

5. D.J. Cosgrove, Inositol Phosphates: **(1980)** Their Chemistry, Biochemistry and Physiology, Elsevier, Amsterdam, , p. 191.

6. D.J. Morre´, W.F. Boss, F.A. Loewus (Eds.), **(1990)** Inositol Metabolism in Plants, Wiley–Liss, **393** p, New York,.

7. Diehl RE, Whiting P, Potter J, Gee N, Ragan CI, Linemeyer D, Schoepfer R, Bennett C, Dixon AF. **(1990)**. Cloning and expression of bovine brain inositol monophosphatase. Journal of Biological Chemistry **256**, 5946–5949.

8. F.A. Loewus, M.W. Loewus, **(1983)** *myo*-Inositol: Its biosynthesis and metabolism, Annu. Rev. Plant Physiol. **34** 137–161.

9. Gillaspy GE, Keddie JS, Oda K, Gruissem W. **(1995).** Plant inositol monophosphatase is a lithium-sensitive enzyme encoded by a multigene family. The Plant Cell **7**, 2175–2185.

10. H. Bohnert, R.G. Jensen, **(1996)** Strategies for engineering water-stress tolerance in plants, Trends Biotechnol. **14** 89–97.

11. Hipps, P. P., Sehgal, R. K., Holland, W. H. & Sherman, W. R. **(1973)**. Identification and partial characterization of inositol: NAD+ epimerase and inosose: NAD(P)H reductase from the fat body of the American cockroach *Periplaneta americana*. *Biochemistry* **12**, 4705–4712

12. Jukanti, A.K., Gaur, P.M., Gowda, C.L. & Chibbar, R.N. **(2012)** Nutritional quality and health benefits of chickpea (*Cicer arietinum* L.): a review. *Br. J. Nutr.* **108**, S11–S26

13. Jun Li,Xinbin Dai,Tingsong Liu and Patrick Xuechun Zhao. **(2012)** LegumeIP: an integrative database for comparative genomics and transcriptomics of model legumes. Nucleic Acids Research Jan ;40(Database issue):D1221-D1229

14. Kumar S, Stecher G and Tamura K **(2016)** MEGA7:Molecular Evolutionary Genetics Analysis verion 7.0 for bigger datasets. Molecular Biology and Evolution **33**:1870-1874

15. Laing WA, Bulley S, Wright M, Cooney J, Jensen D, Barraclough D, MacRae E. (**2004)**. A highly specific l-galactose-1-phosphate phosphatase on the path to ascorbate biosynthesis. Proceeding of the National Academy of Sciences, USA **101**, 16976–16981.

16. M. Horbowicz, R.L. Obendorf, **(1994)** Seed desiccation tolerance and storability: Dependence on flatulence-producing oligosaccharides and cyclitols (review and survey, Seed Sci. Res. **4** 385–405.

17. Neuwald AF, Krishnan BR, Brikun I, Kulakauskas S, Suziedelis K, Tomcsanyi K, Leyh TS, Berg DE. **(1992).** cysQ, a gene needed for cysteine synthesis in Escherichia coli K-12 only during aerobic growth. Journal of Bacteriology **174**, 415–425.

18. Nunez, L. R. & Henry, S. A. **(2006)** Regulation of 1d-*myo*-inositol-3-phosphate synthase in yeast. *Subcell. Biochem.* **39**, 135–156

19. P.P.N. Murthy, Inositol phosphates and their metabolism in plants, in: B.B. Biswas, S. Biswas (Eds.), *myo*-Inositol Phosphates, Phosphoinositides, and Signal Transduction, Subcellular Biochemistry Series, **(1996)** vol. **26**, Plenum Press, New York, , pp. 227–255.

20. Parthasarathy R, Parthasarathy L, Vadnal R**. (1997).** Brain inositol monophosphatase identified as a galactose 1-phosphatase. Brain Research **778**, 99–106.

21. Posternak, T. *The Cyclitols* (Editions Scientifiques Hermann, Paris, **1965**).

22. R.L. Obendorf, **(1997)** Oligosaccharides and galactosyl cyclitolsin seed desiccation (review update), Seed Sci. Res. **7** 63–74.

23. Stec B, Yang H, Johnson KA, Chen L, Roberts MF. **(2000).** MJ0109 is an enzyme that is both an inositol monophosphatase and the 'missing' archaeal fructose-1,6-bisphosphatase. Nature Structural Biology **7**, 1046–1050.

24. Styer JC, Keddie J, Spence J, Gillaspy GE. **(2004).** Genomic organization and regulation of the LeIMP-1 and LeIMP-2 genes encoding myo-inositol monophosphatase in tomato. Gene **326**, 35–41.

25. Suzuki M, Tanaka K, Kuwano M, Yoshida KT**. (2007).** Expression pattern of inositol phosphate-related enzymes in rice (Oryza sativa L.): implications for the phytic acid biosynthetic pathways. Gene **405**, 55–64.

26. T. Peterbauer, A. Richter, **(1998)** Galactosylononitol and stachyose synthesis in seeds of Adzuki bean, Plant Physiol. **117** 165–172.

27. T. Peterbauer, M. Puschenreiter, A. Richter, **(1998)** Metabolism of galactosylononitol in seeds of *Vigna umbellata*, Plant Cell Physiol. **39** 334–341**.**

28. Torabinejad J, Donahue JL, Gunesekera BN, Allen-Daniels MJ, Gillaspy GE. **(2009)**. VTC4 is a bifunctional enzyme that affects myoinositol and ascorbate biosynthesis in plants. Plant Physiology **150**, 951–96.

APPENDIX-A

Medium Composition

1. Luria-Bertani Medium

Composition of LB medium

Composition	gm/l
Bacto tryptone	10
Bacto yeast extracts	5
NaCl	10

For liquid LB media, 25gm of Luria Bertani medium was added to 1L deionised water and mixed to dissolve. The pH was adjusted to 7.0 with 10N NaOH. Medium was sterilised by autoclaving for 20 minutes at 15lbs. For solid media, 1.4% agar was additionally added to liquid media just before autoclaving.

2. MS medium

Full strength or half strength MS media was prepared as per requirement and Sigma MS medium.

Composition	gm/l
MS-Medium	4.4
MES	0.5
Sucrose	10

All the components were dissolved in deionized water. pH of the medium was adjusted to 5.6 with 10 N KOH. For solid media, 0.8% agar was added to the media. Media was sterilized by autoclaving for 20 minutes at 15 lb/sq in. on liquid cycle.

APPENDIX-B

Microbial Strains

E.coli DH5α strain was used for cloning.

E.coli BL21 (DE3) and *E.coli* Lemo21 was used for protein expression.

Agrobacterium tumefaciens GV3101 was used for plant transformation.

Vectors

pJET 1.2 blunt end vector, Bacterial expression vector pET28a, Plant expression vector pCAMBIA1301 were used in this study.

APPENDIX-C

Solutions and buffers

1) IPTG (1 M) – For 1M stock, 1.2 g of IPTG was dissolved in 5 ml of autoclaved sterile MQ water and stored at -20°C till further use.

2) Ethidium Bromide (10 mg/ml) – 100 mg of ethidium bromide was dissolved in 10 ml of water. The tube was wrapped in aluminium foil and stored at room temperature.

3) EDTA (0.5 M; pH - 8.0) – For 0.5 M stock, 93.05 g of disodium EDTA was added to 400 ml of MQ water. pH was adjusted to 8 with NaOH pellets and volume was made up to 500 ml. Solution was sterilized by autoclaving.

4) 50X TAE (1 L) – 242 g Tris base, 57.1 g of glacial acetic acid, 100 ml of 0.5 M EDTA; pH 8.0. Volume was made up to 1L.

5) 10 X DNA gel loading buffer – 0.25 % Bromophenol blue, 50 % (v/v) glycerol in MQ water.

6) 1 M Tris-Cl (pH – 7.5) – For 1 M stock, 157.56g of Tris-Cl was added to 800 ml of water, pH was adjusted to 7.5 with HCl. The volume was made up to 1L.

7) Wash Buffer- For 100ml, 5ml of (50mM Tris-HCl pH 7.5),0.8766g of 150Mm NaCl, 2Mm β-mercaptoethanol, PMSF in 100ml water

8) Lysis Buffer- Wash buffer containing 1% Triton X-10 and 1M urea per gram cell weight.

9) Solubilisation buffer- 50mM Tris-HCl pH 7.5, 8M Urea, 1mM DTT, 1mM PMSF.

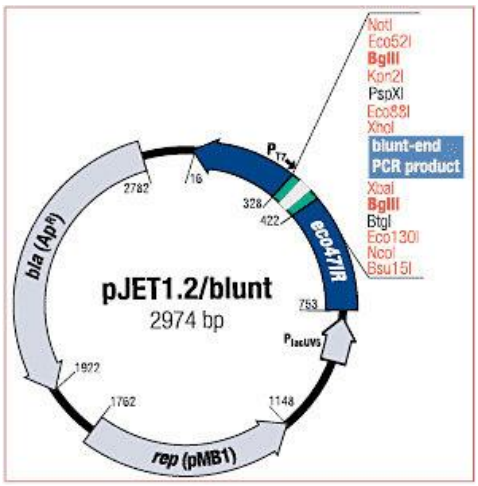
10) Binding Buffer- 20mM Tris HCl pH 7.4, 0.5 NaCl, 20Mm imidazole.

APPENDIX-D

Vector Maps

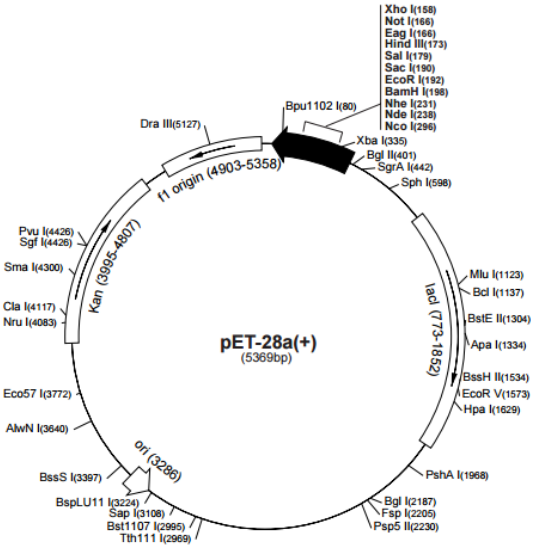
pJET 1.2

This vector consists of a lethal restriction enzyme gene that is disrupted by ligation of a DNA insert into the cloning site. As result, only bacterial cells with recombinant plasmids are able to form colonies. It also has ampicillin resistance gene for bacterial selection.



pET28a vector

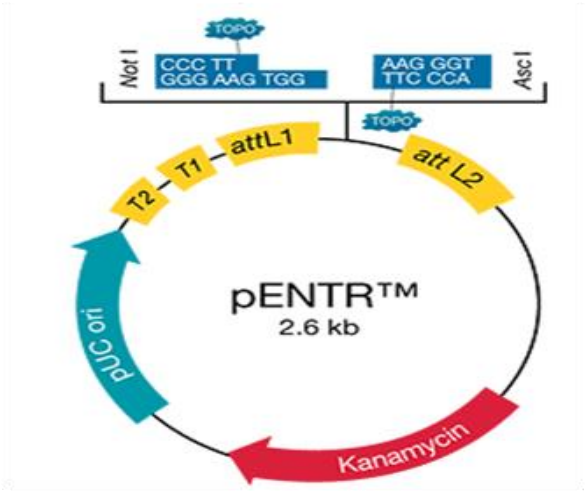
This vector consists of kanamycin antibiotic resistance gene for bacterial selection and lacI gene from the lac operon that code for lac repressor. In order to prevent leaky expression of the gene of interest, T7 promoter is incorporated that only matches RNA polymerase from T7 bacteriophage.



Gateway cloning vectors

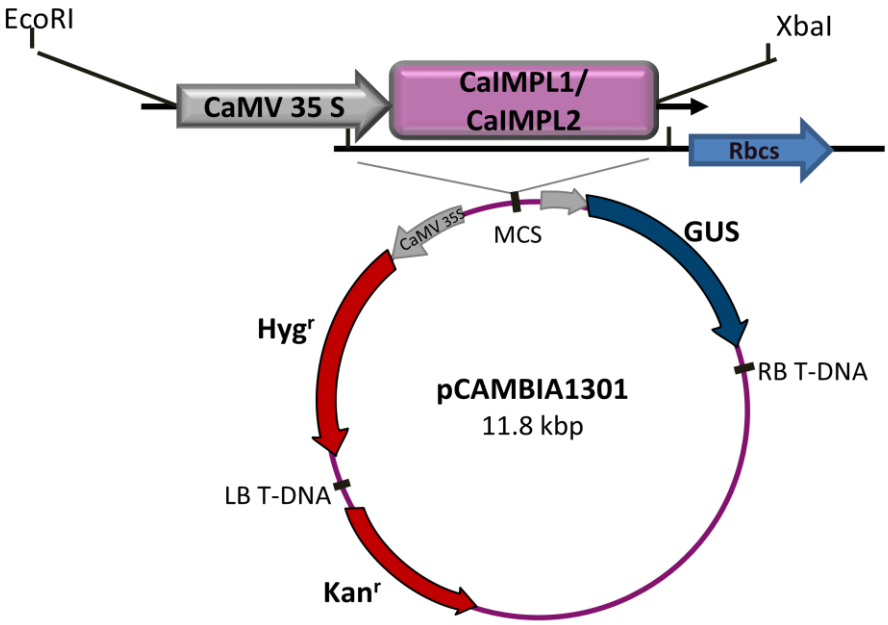
Comprises of a entry and destination vector.

pENTR-D-TOPO is used as entry vector which has kanamycin resistance gene for bacterial selection.



pCAMBIA1301 vector

It is an Agrobacterium binary vector for plant transformation, with hygromycin- and kanamycin-resistance and GUS genes.



APPENDIX-E

Primers used in the project

CaIMPL1(F)	CACCATGTCAATTGTATTCTCCGCAGC	For Gateway-cloning
CaIMPL1(R)	AACGTCTGCCCTGTAGTCCTC	
CaIMPL2(F)	CACCATGTTGTCACAGTGCCATCT	
CaIMPL2(R)	CCGTTGTAATGAATCTAGAGC	
CaIMPL1(F)	GGTACCATGGGAAGGTCTCTAATATT	For cloning into pCAMBIA1301 vector
CaIMPL1(R)	GTCGACATTATCATACAGAGCTTTAA	
CaIMPL2(F)	TCTAGAATGTTAGCTCAGTCGCACTT	
CaIMPL2(R)	GTCGACAGTCACTTGAGTGGCATTGA	

APPENDIX-F

Sequences used in phylogenetic analyses

IMPL1 sequences

>Soybean IMPL1
MSIVFSAAATLSPRSRLLFPPLSFSHTYPTKFHRNGWHPSLSKSRTQPLITNSLLSDKFPTVGSLSTGPIPSTHL
IEVVKTAAQTGAQVVMEAVNKPRNITYKGLTDLVTETDKMSEAAILEVVKKNFEDHLILGEEGGVIGDAASDYLW
CIDPLDGTTNFAHGYPSFAVSVGVLYRGNPAAAAVVEFVGGSMCWNTRLFSATAGGGAFCNGQRIQVSATNQVER
SLLVTGFGYDHDDAWATNIDLFKEFTDVSRGVRRLGAAAVDMCHVALGIVEAYWEYRLKPWDMAAGVLMVEEAGG
TVSRMDGGKFCVFDRSVLVSNGQLHAKLLERIGPATEKLKNKGIDFSLWYKPENYRADV

>Bean IMPL1
MGRSLVFSTNIPLEFSQKPRSFSLLNHSQLCFPQRFIENSQSGYKKIQLLNLKLARNVCTKAALSEITNERKYPK
VGAPSTGPISANQLIQVVETAAKTGAEVVMDAVNKPRNITYKGLTDLVTDTDKMSEAAILEVVRRNFGDHLILGE
EGGIIGDTLSDYLWCIDPLDGTTNFAHGYPSFAVSVGVLFRGNPAAAAVVEFVGGPMAWNTRTFTAIAGGGAFCN
GQKIHASQTDRVEQSLLVTGFGEHDDPWATNIELFKEFTDVSRGVRRLGAAAVDMCHVALGIVEAYWEYRLKPWD
MAAGVLIVEEAGGTVSCMDGGKFCVFDRSVLVSNGVLHAKLLERIAPATEKLKSKGIDFSLWYKPENYRTDL

>Alalfa IMPL1
MMSIVFSTSAAATKLFPPTTYRLQSQTSRSWRLKSSTSCKNSLKSEEKLYSRVGALSTGPVQPAILLEVATTAAQ
TGAKVVMDAVNKPRSITYKGLTDLVTETDKMSEAAILEVVKKNFEDHLILGEEGGIIGDVASDYLWCIDPLDGTT
NFAHGYPSFAVSVGVLYRGNPTAATVVEFVGGPM
CWNTRIFTATAGGGAFCNGQRIEVSVTDQVERSLLVTGFGYEHDEAWATNIELFKEFTDVSRGVRRLGAAAVDMC
HVALGIVEAYWEYRLKPWDMAAGVLMVEEAGGTVSRMDGGKFCVFDRSVLVSNGVIHAKLLERIGPATEGLKSKG
IDFSLWYKPEDYRADV

>Castor bean IMPL1
MGTSLLFSTNISLNFSRLPTSLSPPSYPNQCLPQRFNANSRCGYASINFNTKSIRRNLCPNAVLSETPRNQRQYA
RVGAQSTGPVPLSQLIQVVETAAKTGAEVVMDAVNKPRNISYKGLTDLVTDTDKMSEAAILEVVRKNFGNHLILG
EEGGIIGDTLSDYLWCIDPLDGTTNFAHGYPSFAVSVGVLFRGNPAAAAVVEFVGGPMAWNTRIFSATAGGGAFC
NGQRIHVSQTDKVEQSLLVTGFGYEHDDAWATNIELFKEYTDVSRGVRRLGAAAVDMCHVALGIVEAYWEYRLKP
WDMAAGVLIVEEAGGTVTRMDGGKFCVFDRSALVSNGVLHSKLLERIGPATEKLKSKGIDFSLWFKPENYHTDF

>Chickpea IMPL1
MSIVFSAASNLSWHKDCRQSSPPIGSWRLKSRIQSCKNSLQSDIYTQHRVGARSTGPIQPTHLIQVATTAAQTGA
QVVMDAVNKPRNITYKGLTDLVTETDKMSEAAILEVVKKNFDDHLILGEEGGIIGEAASDYLWCIDPLDGTTNFA
HGYPSFAVSVGVLYRGNPAAATVVEFVGGPMCWNTRIFTATAGGGAFCNGQRIHVSATNQVEQSLLVTGFGYEHD
EAWATNIELFKEFTDVSRGVRRLGAAAVDMCHVALGIVEAYWEYRLKPWDMAAGVLMVEEAGGTVSRMDGGKFCV
FDRSVLVSNGVLHTELLERIGPATEELKSKGIDFSLWYKPEDYRADV

>Arabidopsis IMPL1
MGRSLIFSGNMSLRISHLPRSSLPLQNPISGRTVNRTFRYRCTRILSNSFKSTTRLQTKAVLSEVSDQTRYPRIG
AKTTGTISPAHLLEVVELAAKTGAEVVMEAVNKPRNITYKGLSDLVTDTDKASEAAILEVVKKNFSDHLILGEEG
GIIGDSSSDYLWCIDPLDGTTNFAHGYPSFAVSVGVLYRGNPAAASVVEFVGGPMCWNTRTFSATAGGGALCNGQ
KIHVSKTDAVERALLITGFGYEHDDAWSTNMELFKEFTDVSRGVRRLGAAAVDMCHVALGIAESYWEYRLKPWDM
AAGVLIVEEAGGAVTRMDGGKFSVFDRSVLVSNGVLHPKLLERIAPATENLKSKGIDFSLWFKPEDYHTEL

>Alalfa IMPL2
MLLSQCHLLHSKIPNIEKNNPFRNQYQLQFQPNKLTSPLLLSSSSPKFRIRAMSSSSSPPHQLNHFSDVANKAAN
AAGDVIRKYFRKNNFDIIHKNDLSPVTIADQSAEEAMVSVILDNFPSHAVYGEEKGWRCKQDSADYVWVLDPIDG
TKSFITGKPLFGTLIALLQNGTPILGIIDQPVLKERWIGITGKRTTLNGQEVSTRTCADLSQAYLYTTSPHLFSG
DAEEAFIRVRDKVKIPLYGCDCYAYALLSSGFVDLVVESGLKPYDFLALIPVIEGSGGVITDWKGHQLRWEASPL
SIATSIFLFNISHLKSCLIRNSFLALYIKI

>Soybean IMPL2
MFSQCHFLSHSPIPNTTFRLRAMAPHSTPLELNRFAEVGNKVADAAGEVIRKYFRKNFDVIHKHDLSPVTIADQS
AEEAMVSIILDNFPSHAIYGEENGWRCEEKNADYVWVLDPIDGTKSFITGKPVFGTLVALLQNGTPILGIIDQPV
LRERWIGIAGKRTSLNGQEISTRTCADLSQAYLYTTSPHLFSGDAEEAFIRVRSKVKIPLYGCDCYAYALLSSGF
VDLVVESGLKPYDFLALIPVIEGAGGVITDWKGDKLFWEASPLSIATTLSYLHVAGFNVVAAGDKQIHQQALDSL
QWK

>Bean IMPL2
MFSQCHFLSHSHSPIPNSSLSPTFRLRAMAPHTTPLQLNRFVEVGNTAADAAGEVIRKYFRKNFDVLHKHDLSPV
TIADRSAEEAMVSIILDNFPSHAIYGEENGWRCKEKTADYVWILDPIDGTKSFITGKPVFGTLIALLQNGTPILG
IIDQPVLKERWIGISGKRSTLNGQEISTRTCADLSQAYLYTTSPHLFSGDAEEAFIRVRSKVKIPLYGCDCYAYA
LLSSGFVDLVVESGLKPYDFLALVPVIEGAGGVITDWKGDKLFWEASPLSIATNFNVVAAGDKQIHQQALDSLQW
K

>Castor bean IMPL2
MLSQSQSYFLSQIPKFPLSSTPHFSLNSSTGGSGSGGRVGGTATAVSFQSSGVESPISLKLQTKKLASAMTSNSK
LSDHTDTLHSLVTDGELDRFADVANKVADASGEVIRTYFRKKFDILDKEDSSPVTIADKAAEESMVKIILENFPS
HAIYGEENGWRCKENFSDYVWVLDPIDGTKSFITGKPLFGTLIALLHRGKPILGIIDQPVLKERWIGITGRRTTL
NGEELSTRSCAKLSQAYLYTTSPHLFNGEADEAFTRVRSKVKVPLYGCDCYAYALLASGYVDLVIESGLKPYDFL
SLVPVIEGAGGIITDWKGHHLCWDASPNSRATSFNVLAAGDEQIHQQALDSLEWH

>Arabidopsis IMPL2
MRLHNRDLSIQTFSVVDDWLFRELPTHSKMLAQSHFFSKSFDLIPPQSPALRSANPSLRISSSYSNSRLSFLSSS
AIAVPVSRRRFCLTMASNSKRPNISNESPSELSDTELDRFAAVGNALADASGEVIRKYFRKKFDIVDKDDMSPVT
IADQMAEEAMVSIIFQNLPSHAIYGEEKGWRCKEESADYVWVLDPIDGTKSFITGKPVFGTLIALLYKGKPILGL
IDQPILKERWIGMNGRRTKLNGEDISTRSCPKLSQAYLYTTSPHLFSEEAEKAYSRVRDKVKVPLYGCDCYAYAL
LASGFVDLVIESGLKPYDFLALVPVIEGAGGTITDWTGKRFLWEASSSAVATSFNVVAAGDSDIHQQALESLEWH

>Chickpea IMPL2
MLSQCHLHCYSNNLSIRSPKLRLRAMSSSSSPHQFNHFADVANKAADAAGDVIRKYFRKNFDIIHKHDLSPVTIA
DQTAEEAMVSIILDNFPSHAVYGEEKGWRCRQDSADYVWVLDPIDGTKSFITGKPLFGTLIALLQNGTPILGIID
QPVLRERWIGMTGKRTTLNGQEVSTRTCADLSQAYLYTTSPHLFSGDAEEAFIRVRDKVKIPLYGCDCYAYALLS
SGFVDLVVESGLKPYDFLALVPVIEGSGGVITDWEGHQLRWEASPLSIAISFNVVAAGDKQIHQQALDSLQR

>E.coli SuhB IMPL
MHPMLNIAVRAARKAGNLIAKNYETPDAVEASQKGSNDFVTNVDKAAEAVIIDTIRKSYPQHTIITEESGELEGT
DQDVQWVIDPLDGTTNFIKRLPHFAVSIAVRIKGRTEVAVVYDPMRNELFTATRGQGAQLNGYRLRGSTARDLDG
TILATGFPFKAKQYATTYINIVGKLFNECADFRRTGSAALDLAYVAAGRVDGFFEIGLRPWDFAAGELLVREAGG
IVSDFTGGHNYMLTGNIVAGNPRVVKAMLANMRDELSDALKR

>E.coli CysQ
MLDQVCQLARNAGDAIMQVYDGTKPMDVVSKADNSPVTAADIAAHTVIMDGLRTLTPEIPVLSEEDPPGWEVRQH
WQRYWLVDPLDGTKEFIKRNGEFTVNIALIDHGKPILGVVYAPVMNVMYSAAEGKAWKEECGVCKQIQVRDARPP
LVVISRSHADAELKEYLQQLGEHQTTSIGSSLKFCLVAEGQAQLYPRFGPTNIWDTAAGHAVAAAAGAHVHDWQG
KPLDYTPRESFLNPGFRVSIY

>Human IMPA1
MADPWQECMDYAVTLARQAGEVVCEAIKNEMNVMLKSSPVDLVTATDQKVEKMLISSIKEKYPSHSFIGEESVAA
GEKSILTDNPTWIIDPIDGTTNFVHRFPFVAVSIGFAVNKKIEFGVVYSCVEGKMYTARKGKGAFCNGQKLQVSQ
QEDITKSLLVTELGSSRTPETVRMVLSNMEKLFCIPVHGIRSVGTAAVNMCLVATGGADAYYEMGIHCWDVAGAG
IIVTEAGGVLMDVTGGPFDLMSRRVIAANNRILAERIAKEIQVIPLQRDDED

>Human IMPA2
MKPSGEDQAALAAGPWEECFQAAVQLALRAGQIIRKALTEEKRVSTKTSAADLVTETDHLVEDLIISELRERFPS
HRFIAEEAAASGAKCVLTHSPTWIIDPIDGTCNFVHRFPTVAVSIGFAVRQELEFGVIYHCTEERLYTGRRGRGA
FCNGQRLRVSGETDLSKALVLTEIGPKRDPATLKLFLSNMERLLHAKAHGVRVIGSSTLALCHLASGAADAYYQF
GLHCWDLAAATVIIREAGGIVIDTSGGPLDLMACRVVAASTREMAMLIAQALQTI
NYGRDDEK

>Yeast IMP1
MTIDLASIEKFLCELATEKVGPIIKSKSGTQKDYDLKTGSRSVDIVTAIDKQVEKLIWESVKTQYPTFKFIGEES
YVKGETVITDDPTFIIDPIDGTTNFVHDFPFSCTSLGLTVNKEPVVGVIYNPHINLLVSASKGNGMRVNNKDYDY
KSKLESMGSLILNKSVVALQPGSAREGKNFQTKMATYEKLLSCDYGFVHGFRNLGSSAMTMAYIAMGYLDSYWDG
GCYSWDVCAGWCILKEVGGRVVGANPGEWSIDVDNRTYLAVRGTINNESDEQTKYITDFWNCVDGHLKYD

>Yeast IMP2
MVLTRQVLEEVENTFIELLRSKIGPLVKSHAGTNFCSYDDKANGVDLVTALDKQIESIIKENLTAKYPSFKFIGE
ETYVKGVTKITNGPTFIVDPIDGTTNFIHGYPYSCTSLGLAEMGKPVVGVVFNPHLNQLFHASKGNGAFLNDQEI
KVSKRPLILQKSLIALEGGSERTEGSQGNFDKKMNTYKNLLSESGAFVHGFRSAGSAAMNICYVASGMLDAYWEG
GCWAWDVCAGWCILEEAGGIMVGGNCGEWNIPLDRRCYLAIRGGCESMEQKRFAESFWPHVAGELEY

>Halobacterium SuhB
MHDTDRVSVAERAARTGGAVALDAFRTGIDVETKSGKTDVVTEADRTAQRRVVDVIDAVYDEDTIVGEEADALKT
VPDAGAAWVIDPIDGTNNFVRDTQLWATAVAAVVDGTPVAACNRFPALEDTYIAGADGATLNGTPISVSEKTDPE
TFVVAPTIWWDFDRRDEYAAACAAIVERFGDMRRYGCAQAVLSMVASGQLEATITNVVANPWDSVAGVHLVRQAG
GVVTDIDGTRWTPGATGLVASNGTAHDAVLAAAQEIRAAAE