EVOLUTIONARY DEVELOPMENT AND FUNCTIONAL ROLE OF PLANT NATRIURETIC PEPTIDE (PNP)-B

RUNYARARO MEMORY HOVE

SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE (MSc) - BIOCHEMISTRY

IN THE

FACULTY OF SCIENCE & AGRICULTURE

AT THE

UNIVERSITY OF FORT HARE

University of Fort Hare
Together in Excellence

FEBRUARY 2009

SUPERVISOR: PROF. GRAEME BRADLEY

ABSTRACT

Plant natriuretic peptides (PNP) are novel peptides which, like in vertebrates, have been shown to have a function associated with water and salt homeostasis. Two PNP-encoding genes have been identified and isolated from *Arabidopsis thaliana,* namely; AtPNP-A and AtPNP-B. In this study, the focus was on PNP-B, which has not been extensively studied. Bioinformatic analysis was done on the AtPNP-B gene. This included the bioinformatic study of its primary structure, secondary structure, tertiary structure, transcription factor binding sites (TFBS) and its relation to other known proteins. The AtPNP-B gene was shown to be a 510 bp long, including a predicted 138 bp intron. AtPNP-B was also shown to have some sequence similarity with AtPNP-A and CjBAp12. The TFBS for AtPNP-B and OsJPNP-B were compared and they comprised of TFBS that are related to water homeostasis and pathogenesis. This suggested two possible functions; water stress and homeostasis and a pathogenesis related function for PNP-B. Following bioinformatic analysis, the heterologous expression of the AtPNP-B was attempted to investigate whether the AtPNP-B gene encoded a functional protein and to determine the functional role of PNP-B. However, expression was unsuccessful. An evolutionary study was then carried out which revealed that there were some plants without the intron such as, rice, leafy spurge, oilseed rape, onion, poplar, sugar cane, sunflower and tobacco. These plants would therefore be used for expression and functional studies in the future. The evolutionary studies also revealed that PNP-B had a relationship with expansins and the endoglucanase family 45. Other PNP-B related molecules were also obtained from other plant genomes and therefore used in the construction of a phylogenetic tree. The phylogenetic tree revealed that AtPNP-B clustered in the same group as CjBAp12 while AtPNP-A had its own cluster group. There were also other PNP-B like molecules that clustered in the same group as expansins (α- and β-). Thus, we postulate that, like PNP-A, PNP-B also has a possible function in water and salt homeostasis. However, due to the clustering

of AtPNP-B into the same group as CjBAp12, a possible role of PNP-B in pathogenesis-related response is also postulated.

Keywords: plant natriuretic peptides, evolutionary development, phylogenetic tree, functional role

DECLARATION

I, **Runyararo Memory Hove** declare that the dissertation entitled 'Evolutionary development and functional role of PNP (Plant Natriuretic Peptide) - B' submitted for the degree of MSc (Biochemistry) at the University of Fort Hare is my own work and has not been submitted for any degree or examination at any other university. I further declare that all sources of my information have been quoted as indicated in the text and/or list of reference.

_____

Runyararo Memory Hove

DEDICATION

This study is dedicated to my parents, ‘Baba naMai Rue’.

ACKNOWLEDGEMENTS

I would like to thank the Lord God Almighty for the guidance, wisdom and the strength to carry out this research till completion successfully.

I am greatly indebted to my mentor and supervisor, Professor Graeme Bradley for his expert advice and guidance during this research. I am sincerely grateful to him for accommodating my shortcomings and encouraging me to persevere until I completed my research project. Thank you for helping to turn me from a ‘dependent’ person to an ‘independent’ researcher.

My sincere gratitude also goes to my family for always supporting my decisions and providing moral support, unconditional love and encouragement. I would also like to extend a great thank you to my fiancé, Munyaradzi Wellington Nyadzayo. Thank you for the emotional support my love and for being the greatest graphics designer.

Thank you to the ‘Plant Stress Group’ for the support, advice and encouragement.

I would like to thank the National Research Foundation for the financial support in my studies and the GMRDC for funding the research.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AGE – Agarose Gel Electrophoresis

ANP – Atrial Natriuretic Peptide

BLAST –Basic Local Alignment Search Tool

BNP – Brain Natriuretic Peptide

Bp – Base pairs

$CaCl_2$ – Calcium chloride

cGMP – Cyclic Guanosine Monophosphate

CNP – C-type Natriuretic Peptide

ºC – Degrees Celsius

DNA – Deoxyribonucleic Acid

EXP – alpha Expansin

EXPB – beta Expansin

HR – Hypersensitivity Response

IAA – Indole-3-acetic Acid

IDA – Abscission gene

INA - 2,6-dichloro-isonicotinic acid

IPTG – Isopropyl-β-D-thiogalactopyranoside

irPNP – immunoreactant plant natriuretic peptide

ISR - Induced systemic resistance

LB – Luria Bertani

LRR-RLK – Leucine-rich Repeat Receptor-like Kinase

LTP - Lipid transfer proteins

LY 83583 – 6-anilino-5, 8- quinolinequinone

kDa – KiloDaltons

MALDI TOF MS - Matrix-assisted laser desorption ionization time-of-flight mass spectrometry

µM – Micromolar

µL – Microlitre

µg – Microgram

mg – Milligram

mM – Millimolar

mL – Millilitres

M – Molar

NCR – Nodule-specific Cysteine-rich

Ni-NTA – Nickel-Nitrilotriacetic acid

NMR – Nuclear Magnetic Resonance

PCR – Polymerase Chain Reaction

PDB – Protein data bank

PNP – Plant Natriuretic Peptide

PIs – Protease inhibitors

PR – Pathogen related

RALF – Rapid Alkalinization Factor

RBS – Ribosomal binding site

RNA – Ribonucleic Acid

ROS – Reactive Oxygen Species

RT-PCR – Reverse Transcriptase Polymerase Chain Reaction

SAR - Systemic acquired resistance

SCR - S-locus cysteine-rich

SD – Shine-Dalgarno

SDS – Sodium Dodecyl Sulphate

SDS-PAGE – Sodium Dodecyl Sulphate Polyacylamide Gel Electrophoresis

SP – Signal peptide

TEMED – N, N, N, N – Tetramethylethylenediamine

TFs – Transcription factors

TFBS – Transcription Factor Binding Sites

TIRs - Translation initiation regions

TL - Thaumatin-like

TSS – Transcription start site

VNP- Ventricular Natriuretic Peptide

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF TABLES

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER ONE: LITERATURE REVIEW

1. Introduction

For many years, the insight into intercellular signalling in plants was based upon the knowledge of the five classical hormones namely; auxin, cytokinin, ethylene, gibberellin, and abscisic acid. However, biochemical and genetic studies have identified molecules that play crucial roles in plant growth and development, including defence mechanisms in response to wounding by pests, the control of cell division and expansion and pollen self-incompatibility (Lindsey *et al*., 2002). These range from nutrients such as nitrate and glucose through jasmonate, salicylate and larger molecules such as brassinosteroids and nodulation factors (Gehring, 1999).

There is evidence that signal molecules, once considered to be specific to animal systems, are also found and function in plants such as steroids, sterols and lipid derivatives (Bishop & Yokota, 2001). A natriuretic peptide system was discovered in plants, when radioimmunoassays were performed on tissue extracts from florida beauty (*Dracena godseffiana*) (Vesely & Giordano, 1991). Further studies showed that NP-like proteins were present in several other plant species and data from high performance gel permeation chromatography projected that pro-ANPs and NPs from plants share similar molecular masses with vertebrate ANPs (Vesely *et al.,* 1993).

The discovery of this novel protein peptide system in plants led to a more in-depth study of the plant hormones already known (‘classical and ‘non-classical’) and the link between these hormones and the newly discovered plant natriuretic peptide system.

1.1 Classical Hormones

1.1.1 Auxin

Auxin is a collective term for a group of hormones that has varying effects on plants and they were the first plant hormones to be discovered. One example of an auxin includes indole-3- acetic acid (IAA), which is the main chemical form of auxins and is involved in many physiological processes in plants (Arteca, 1996). Others include indole-3-butyric acid (IBA), phenyl acetic acid and 4-chloro - IAA. Auxins are responsible for promoting cell elongation, a process that is required before differentiation of a cell. They also promote growth by stimulating the formation of lateral promordia (Kende & Zeevart, 1997). Auxins are also involved in cell division by increasing levels of $p34^{cdc2}$-like proteins in cultured tobacco cells (John *et al*., 1993). IAA is synthesized from typtophan through several different routes (Bandurski *et al*., 1995). To date, three pathways for the synthesis of IAA have been proposed, namely via indole-3-pyruvic acid, tryptamine or indole-3-acetonitrile (Kende & Zeevart, 1997). IAA can also occur through a non-typtophan pathway (Normanly *et al.,* 1995).

1.1.2 Gibberellins (GAs)

Gibberellins are growth promoters and have a similar effect as auxins, in that they promote cell division and elongation, break seed dormancy and seed germination. Gibberellins consist of a large family of tetracyclic compounds, of which to date more than 100 Gibberellic Acids (GA) have been identified (Hisamatsu *et al*., 1998). GA_3 or gibberellic acid is the most common form and originally isolated from the fungus *Gibberella fujikuroi* and is known to have a high activity in promoting growth of dwarf plants, dormant buds and dormant seed (Sponsel, 1995). The precursor of gibberellins is mevalonic acid (Macmillan, 1997). Gibberellic acid can promote the

growth of a genetic deficient plant to full size, where the genotype is for minimal growth (Kende & Zeevart, 1997).

1.1.3 Cytokinins

Unlike other hormones, cytokinins are found in both plants and animals. Several bacteria, including Agrobacterium, also produce cytokinins (Gaudin *et al*., 1994). Naturally occurring cytokinins are synthesized from purine. The first naturally occurring cytokinin was extracted from immature kernels of *Zea Mays* and named zeatin or 6-(4-hydroxy-3-methylbut-trans-2-enylamino) purine. The most biologically active cytokinin is not found in plants, but was extracted from autoclaved herring sperm DNA and called kinetin or 6-(furfurylamino) purine (McGaw & Burch, 1995). Other common cytokinins are benzyl adenine (BAP) and isopentenyl adenine (IP).

Cytokinins are involved in cell division and work in conjunction with auxin, to increase the level of a $p34^{cdc2}$- like protein (John *et al*., 1993) and also to induce morphogenesis and to inhibit apical dominance. They also promote stomatal opening (Davies, 1995). Cytokinins are also used to delay aging and death, known as senescence.

1.1.4 Ethylene

Ethylene is unique, as compared to other hormones, in that it can only be found in the gaseous form. Ethylene can also be metabolized into ethylene oxide and ethylene glycol. Since ethylene is in the gaseous form it can easily and readily diffuse from plant tissue, making it unnecessary for metabolism to become part of its removal process (Kende & Zeevart, 1997). Plants produce ethylene in response to stress signals and ethylene is often found in high concentrations within cells at the end of the plant's life. Ethylene enhances fruit ripening, causes leaves to droop

(epinasty) and drop (abscission) and promotes senescence (Sponsel, 1995). Ethylene inhibits elongation of terrestrial plants and causes thickening of the stem (Lang *et al*., 1982), but promotes the rapid elongation of semi-aquatic plants upon submergence (Voesenek *et al*., 1992). Ethylene is synthesized from methionine (McKeon *et al.,* 1995).

1.1.5 Abscisic Acid (ABA)

Abscisic acid is a general plant growth inhibitor and high concentrations of abscisic acid (ABA) are present in the guard cells and play a role in the stomatal closing during periods of drought stress (Zeevaart & Creelman, 1988). ABA also induces dormancy, prevents seeds from germinating, promotes abscission (Walton, 1980) and induces storage protein synthesis in seeds (Kende & Zeevart, 1997; Davies, 1995). Carotenoids are precursors for ABA, with xanthoxin as the intermediates (Zeevaart *et al*., 1991).

1.1.5.1 Possible Interaction of ABA with Plant Natriuretic Peptides

ABA is the most likely of the five 'classical' plant hormones to interact with the PNP signalling processes, because it is involved in the regulation of the response to abiotic stresses, such as drought, salinity and cold. ABA has a major role in maintaining plant water status under these abiotic stress conditions by stimulating stomatal closure and regulating the expression of many genes (Finkelstein *et al*., 2002; Xiong *et al*., 2002). ABA signalling is a well studied and complex process. ABA operates through a labyrinth of multiple redundant perception and signalling pathways that in turn are co-regulated by other hormones, environmental stimuli and metabolic status (Schroeder *et al*., 2001; Finkelstein *et al*., 2002; Le´on & Sheen 2003; Razem *et al*., 2006). A noteworthy antagonism is that exogenous ABA promotes stomatal closure (Schroeder *et al*. 2001), whereas PNP stimulates stomatal opening (Pharmawati *et al*., 1998*a*, 2001; Morse *et al*., 2004). ABA levels accumulate in plants in response to abiotic stresses, such as drought and

salinity, as a result of enhanced synthesis of ABA (Xiong *et al*., 2002). Levels of irPNP are also increased in response to saline and osmotic stresses (Rafudeen *et al*., 2003).

1.2 Non-classical Hormones

1.2.1 Oligosaccharins

Oligosaccharins are synthesized in the cell walls and consist of two major groups, namely; pectic oligosaccharins and xyloglucan oligosaccharins. These oligosaccharins play roles in the induction of defense reactions following pathogen attack (Aldington & Fry, 1993), auxin antagonism (Warneck *et al*., 1998) and induction of root nodules. Pectic oligosaccharins display 'hormone-like' activity, contributing to the natural control of plant metabolism, growth and development (Aldington & Fry, 1993). Pectic oligosaccharins also inhibit auxin action and function as elicitors to initiate pathogen defense response (Creelman & Mullet, 1997). Xyloglucan oligosaccharins can inhibit auxin-, gibberellins-, proton- and fussicoccin-induced growth and inhibit endogenously regulated growth at a concentration of 1nM (McDougall & Fry, 1989). Some oligosaccharins can also promote stem elongation at higher concentrations (McDougall & Fry, 1990).

1.2.2 Brassinolides

Brassinolides are natural steroid hormones that were first discovered in Brassica pollen (Grove *et al.*, 1979). Its sources are reproductive organs and growing tissue. It plays a role in stimulation of cell division and elongation (Mussig & Altman, 1999), promotion of xylem differentiation (Oh & Clouse, 1998), activation of H^+ pump and increases the success of fertilization (Khripach *et al*., 2000). To date, more than 40 related steroids, called the brassinosteroids, have been identified in the plant kingdom (Fujioka, 1999; Khripach *et al*., 2000).

1.2.3 Jasmonates (JA)

Jasmonates are volatile fatty acid derivatives. Its source is in young organs and developing reproductive structures. The functions of JA in plants are varied and include inhibition of the germination of non-dormant seeds and the stimulation of the germination of dormant seeds (Bewley, 1997). Experiments by Berger *et al.*, (1996) suggest that JA may stimulate seed germination by decreasing sensitivity to ABA. In addition to its effects on seed germination, JA also strongly inhibits root growth (Berger *et al.*, 1996). JA also promotes senescence, abscission, fruit ripening and pigment formation (Staswick, 1995). It also plays a role in the induction of defense genes and proteins (Bewley, 1997), stimulation of dormant tissue germination (Bergey *et al*., 1996) and inhibition of root growth. Creelman & Mullet, (1997) reviewed several roles of JA action in plant defense. It is observed that there is an accumulation of JA in wounded plants as well as in plant cell culture treated with elicitors of pathogen defense.

1.2.4 Salicylic Acid

Salicylic acid is a phenolic compound, which has its source in flowers and infected organs. Salicylic acid is derived from the shikimate-phenyl propanoid pathway in higher plants (Metraux, 2002). It is a key-signalling component in systemic acquired resistance (SAR); a defence mechanism that confers a broad-spectrum resistance to pathogens (Delaney, 1997; Durner *et al*., 1998). SAR confers resistance to plants by interfering with replication of pathogens and/or inhibiting their systemic movement in the plant (Durner *et al.,* 1998). Usually SAR follows after initial and rapid hypersensitivity response (HR) characterized by responses such as necrotic lesion around pathogens and sustained burst of reactive oxygen species (ROS) (Lamb and Dixon, 1997). Salicylic acid plays a role in the activation of pathogen defense genes (Shah & Klessig, 1999) and induction of flowering (Yalpani *et al*., 1991).

One of the most studied functions of salicylic acid is its association with, and involvement in plant resistance response to different pathogen attacks (Yalpani & Raskin, 1993; Alvarez, 2000). It was also reported that salicylic acid accumulates during exposure to ozone or UV light (Sharma *et al.,* 1996). Salicylic acid also reduces transpiration and reverses ABA-induced stomatal closure (Raskin, 1995).

1.3 Peptide Phytohormones

Protein/peptide signalling molecules are well characterised in animal systems where they create diversity in signalling. However, it is only relatively recently that they have been recognised in plants (Wang *et al.,* 2007). Over the last decade it has become apparent that plants also contain peptidic signalling molecules that play vital roles in cell-to-cell communication (Matsubayashi *et al*., 2001; Lindsey *et al*., 2002). Peptide hormones have been shown to affect development (Topping & Lindsey, 1997), reproduction (Takayama *et al.,* 2000), nodulation (Campalans *et al.,* 2004) and defense (Schaller & Ryan, 1995).

1.3.1 Peptide Phytohormones with Non-defense Signalling Roles

1.3.1.1 Peptide Indole-3-acetic acid

The primary auxin in plants is indole-3-acetic acid (IAA). Although other compounds with auxin activity, such as indole-3-butyric acid (Epstein & Ludwig-Müller, 1993), phenyl acetic acid, and 4-chloro-IAA, are also present in plants (Normanly *et al.,* 1995), little is known about their physiological role. Indole-3-acetic acid (IAA) is involved in virtually all aspects of plant growth and development (Davies, 1995). Plants produce active IAA by *de novo* synthesis and by hydrolyzing IAA conjugates (Normanly *et al.,* 1995; Bartel, 1997; Normanly, 1997).

Indole-3-acetic acid (IAA) occurs not only in the free form but also appears in ester-conjugated forms with glucans (Piskornik & Bandurski, 1972), glucose, inositol (Cohen & Bandurski, 1982) and glycoprotein (Percival & Bandurski, 1976). Amide conjugations to aspartate (Cohen, 1982), glutamate (Epstein *et al.,* 1984) and peptides (Bialek & Cohen, 1986) also exist. IAA conjugation activity is widely distributed in the plant kingdom from mosses to angiosperms (Sztein *et al.,* 1995), and most IAA in plant tissues is conjugated via its carboxyl group (Cohen & Bandurski, 1982; Bandurski *et al.,* 1995).

Different conjugates may perform different functions in the plant. For example, IAA–Asp is an intermediate in IAA destruction (Tsurumi & Wada, 1986; Monteiro *et al.,* 1988; Tuominen *et al.,* 1994). Bialek and Cohen (1986) purified a 3-kDa Phaseolus peptide IAA that may have roles in IAA metabolism, transport and storage (Cohen & Bandurski, 1982). Other conjugates may serve as reservoirs of inactive IAA that can be hydrolyzed to supply the plant with active hormone, as in maize germination when conjugate hydrolysis provides free IAA to the developing seedling (Epstein *et al.,* 1980). Some IAA conjugates appear to serve functions such as IAA storage forms in seeds and hormonal homeostasis. The *iaglu* gene in maize, which encodes an enzyme that esterifies IAA to glucose, has been cloned (Szerszen *et al.,* 1994). In Arabidopsis, a gene family that encodes IAA conjugate hydrolases has been identified (Bartel, 1997).

1.3.1.2 Phytosulphokines (PSKs)

Phytosulphokines are disulphated pentapeptides that have mitogenic activities and can stimulate colony formation of rice protoplasts at concentrations > 10^{-8} M. PSK-alpha [Tyr(SO_3H)-Ile-Tyr(SO_3H)-Thr-Gln] and PSK-beta [Tyr(SO_3H)-Ile-Tyr(SO_3H)-Thr] are two sulphated peptides that have been obtained from *Asparagus officinalis* (Matsubayashi & Sakagami, 1996). The sulphation of the peptide occurs at a tyrosine residue preceded by an acidic residue (Hanai *et al.,*

2000), and PSK-alpha is less abundant and more active than its truncated form, PSK-beta (Yang *et al.*, 1999b).

Binding assays with [^{35}S]PSK-alpha demonstrated that rice suspension cells contain both specific high-(K_d 10^{-9}) and low-affinity (K_d 10^{-7}) saturable PSK binding sites on the surface. Since PSK showed high ligand binding specificity, it was estimated that each cell contains $\geq 10^5$ binding sites for each type. PSK binding is not affected by auxin or cytokinin, suggesting that novel signal transduction pathways might be involved in PSK-dependant activation of gene induction (Matsubayashi *et al*., 1997).

PSK is highly conserved amongst plants and is involved mainly in cell proliferation and differentiation (Matsubayashi & Sakagami, 1996). It has been suggested that PSK helps to initiate proliferation and/or differentiation of mature quiescent tissues (Matsubayashi & Sakagami, 2006). In addition, PSK has many other biological functions, such as the enhancement of chlorophyll synthesis under etiolated or short-day growth conditions (Yamakawa *et al.,* 1999), the promotion of adventitious root (Yamakawa *et al.,* 1998) and bud (Yang *et al.,* 1999a) formation, and stimulation of the differentiation of the tracheary element (Matsubayashi *et al.*, 1999) and somatic embryogenesis (Kobayashi *et al.*, 1999).

1.3.1.3 CLAVATA3 (CLV3) and CLE

CLAVATA3 was first isolated from *Arabidopsis* meristem. CLAVATA3 encodes a 96-amino acid polypeptide with a predicted 18-amino acid secretion signal at the N-terminus (Fletcher *et al.,* 1999). After being C-terminally processed to a 12-amino-acid peptide and hydroxyprolinated (Ito *et al.*, 2006; Kondo *et al.,* 2006; Ni & Clark, 2006), acts as an extracellular ligand. CLV3 has also been reported to be involved in growth coordination in the plant (Schopfer *et al*., 1999).

CLAVATA1 and CLAVATA3 are expressed in the shoot apical meristem, as detected by *in situ* hybridization and they are also involved in shoot meristem organization (Fletcher *et al.*, 1999). It was found that CLAVATA1 and CLAVATA3 act in the same pathway, suggesting that CLAVATA3 may encode CLAVATA1, making them mutually epistatic (Clark *et al*., 1997; Fletcher *et al*., 1999).

CLV3 processing occurs at the conserved residue, arginine-70, at the beginning of the CLE domain. CLE is a conserved sequence located at the COOH-terminus of CLV3 (Ni & Clark, 2006). CLE genes encode small secreted peptides (14 amino acid residues) that are processed from a larger polypeptide and then modified (Matsubayashi & Sakagami, 2006). These peptides have been found in a range of tissues, and mutant analyses of different CLEs have shown that they may act as ligands in different signalling pathways regulating plant development (Hobe *et al.*, 2003; Fiers *et al.*, 2004).

1.3.1.4 POLARIS

POLARIS (PLS) gene expression has been detected in the early stages of embryo and seedling development (Topping & Lindsey, 1997) where it encodes a 36-amino-acid peptide necessary for correct root growth, leaf vascular patterning (Casson *et al.*, 2002) and correct auxin-cytokinin homeostasis (Matsubayashi & Sakagami, 2006). Further studies are required for native peptide isolation. Mutant analysis has shown that PLS expression is repressed by ethylene and induced by auxin. This auxin-ethylene interaction leads to modulation in cell division and expansion, eventually influencing root growth and lateral root development (Chilley *et al.*, 2006).

1.3.1.5 Inflorescence Deficient in Abscission Gene

The inflorescence deficient in abscission (IDA) gene is expressed in the floral organ abscission zone where the floral organs are being shed. *Arabidopsis ida* mutants are deficient in inflorescence abscission (Butenko *et al.,* 2003). IDA over-expressed lines lead to phenotypes such as, the ectopic loss of organs, premature floral organ abscission and the accumulation of arabinogalactan in degraded cell wall middle lamella (Stenvik *et al.,* 2006), whereas *ida* mutants retain floral organs indefinitely. Butenko *et al.,* (2006) produced an IDApromoter::GUS line in the *etr1-1* mutant background (an ethylene-insensitive mutant), and observed a lack of GUS staining in the abscission zone, demonstrating that ethylene plays a role in IDA gene regulation.

IDA encodes a 77-amino-acid peptide, which may be processed to a similar peptide that acts as a ligand. It has been suggested that IDA is recognised by HAESA, a leucine-rich repeat receptor-like kinase (LRR-RLK) found at the base of petioles and pedicles (Jinn *et al.,* 2000). Isolation and characterization of the native IDA peptide may help in the understanding of ligand-receptor interactions (Matsubayashi & Sakagami, 2006) and, eventually, in the engineering of plants that retain their floral organs. This is an important feature in cut flowers and may also affect the timing of pollination and, ultimately, the yield.

1.3.1.6 ROTUNDIFOLIA4/DEVIL1

ROTUNDIFOLIA4 (ROT4) and DEVIL1 (*dvl1*, mutants have a horned fruit phenotype) belong to a 22-member ROT-FOUR-LIKE (RTFL) gene family in Arabidopsis, which encodes small peptides (ROT, 53 amino acids; DVL1, 51 amino acids) (Narita *et al.,* 2004; Wen *et al.,* 2004). These short peptides are involved in the regulation of polar cell proliferation through-out development, and over-expression leads to smaller leaves, shorter bolts and smaller inflorescences. Although their transcripts have been found in various tissues, the actual peptides

have not been isolated or structurally characterized. However, over-expression of a domain encoding a 29-amino-acid region conserved in RTFL family members leads to aberrant phenotypes (Narita *et al.*, 2004).

1.3.1.7 S-locus Cysteine-rich Protein/S-locus Protein 11

S-locus cysteine-rich (SCR) proteins or S-locus protein 11 (SP11) originated from Brassica pollen. It is a 50-53 amino acid protein, synthesized from 74 – 81 amino preprotein (Schopfer *et al*., 1999). Through gain-of-function experiments (Schopfer *et al*., 1999) and a pollination bioassay, it has been confirmed that SCR/SP11 is the key pollen protein that controls self-incompatibility (Takayama *et al.,* 2000), a plant-specific mechanism that prevents self-fertilization and leads to diversity within species. Expression studies reveal that SCR genes are expressed gametophytically in the microspores and in the tapetum and that the proteins might diffuse between developing pollen grains (Schopfer & Nasrallah, 2000).

Mature SP11 homologues generally contain eight conserved cysteine residues (Watanabe *et al.,* 2000). Immunoprecipitation of S8-SP11 purified from pollen under non-reducing conditions, followed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS), revealed the mass value of 5716, suggesting that mature SP11 is a processed monomer with oxidized cysteine residues involved in four intramolecular disulphide bonds (Takayama *et al.,* 2001). Furthermore, Mishima *et al.,* (2003) established the three-dimensional structure of S8-SP11 via nuclear magnetic resonance (NMR) spectroscopy, and demonstrated that its α/β-sandwich structure is similar to that of plant defensins. Through structure-based sequence alignment and homology modeling of allelic SP11, a hyper-variable (HV) region was identified that was suggested to be the specific binding site for the stigma receptor (Mishima *et al.,* 2003).

Orthologues of SCR/SP11 have been found in Arabidopsis in a range of tissues, and are named SCR-related (SCRL; Fletcher *et al.,* 1999). The function of this family of 28 genes, encoding 4.4–9.5 kDa peptides, remains uncharacterized (Matsubayashi & Sakagami, 2006)

1.3.1.8 4-kDa Peptide

Following the discovery of a 43-kDa basic hormone receptor glycoprotein with kinase activity, an extensive search was performed to isolate its hormonal peptide ligand (Watanabe *et al.,* 1994). A peptide with a molecular mass of 3920 Da and with 37 amino acid residues was isolated from radicles of germinated soybean seed, and called leginsulin, insulin-like activity or 4-kDa peptide (Watanabe *et al.,* 1994). Subsequently, it was demonstrated that the peptide co-localizes with and can bind to the α-subunit of the 43-kDa peptide (Shang *et al.,* 2004). Sequence comparison analysis between mRNA and the mature peptide revealed that the 4-kDa peptide is a processed polypeptide from a precursor protein containing a signal peptide that is removed co-translationally, followed by a 4-kDa peptide that is cleaved post-translationally, a linker and a 6-kDa peptide (Watanabe *et al.,* 1994).

The peptide is specific to legumes and the sequence is highly conserved, with six cysteines establishing three disulphide bonds stabilizing a T-knot scaffold, which is essential for ligand function (Watanabe *et al.,* 1994; Yamazaki *et al.,* 2003). MS analysis revealed the presence of two signals with a 56.6 Da mass difference, proposing that the active form has lost a glycine residue from its COOH-terminus, similar to a number of animal peptide hormones (Watanabe *et al.,* 1994). The 4-kDa peptide stimulates protein kinase activity (Watanabe *et al.,* 1994), resulting in an acceleration in callus growth and differentiation at an early developmental stage, and regulates cell proliferation (Yamazaki *et al.,* 2003). Defensive action has been demonstrated for a pea 4-kDa peptide (PA1b) against cereal weevils *Sitophilus spp.* (Gressent *et al.,* 2003).

1.3.1.9 Early Nodulin 40 (ENOD40)

Early nodulin 40 (ENOD40) is a gene with several short ORFs, and therefore it is not clear whether these ORFs are translated (Campalans *et al.*, 2004; Guzzo *et al.*, 2005). Gulyyaev & Roussis (2007), through a study of predicted secondary structures and comparative analysis of plant ENOD40 RNAs, determined the most conserved structural features. These characteristics showed similarity to expansion segments found in structural RNAs [rRNAs, RNase P and signal recognition particle (SRP) RNAs] and higher sequence conservation in ENOD40 RNA molecules than in deduced peptide sequences. These findings suggest that ENOD40 in RNA format is favoured and that its short peptide form only appears in certain plant families.

Early nodulins are peptide signals involved in the early responses of non-legume and legume hosts to nodulation factors and/or early nodulation steps. ENOD40 expression has been detected in the root pericycle and in dividing cortical cells of the nodule primordium in legumes, suggesting a role in the sensitization of cells to division-inducing signals (Charon *et al.*, 1999). The ENOD40 gene is expressed in the pericycle of legume roots early in the nodulation process, adjacent to the protoxylem poles, before the cortical cells divide to form the nodule itself (Yang *et al.*, 1993; Compaan *et al.*, 2001). The implication is that ENOD40 therefore counteracts the effects of ethylene to promote cortical divisions and nodule formation (Lindsey *et al.*, 2002).

ENOD40 is expressed in response to Nod factors and classical hormones such as auxin and cytokinin in legumes and non-legumes (Horvath *et al.*, 1993). ENOD40 has a role in root nodulation and unlike systemin; it is not synthesized via a precursor. ENOD40 is partially responsible for the induction of cortical cell divisions in the early phase of root nodule morphogenesis (Cooper & Long, 1994; Minami *et al.*, 1996). During nodule formation there is a hormonal imbalance in the root cortex which is affected by ENOD40 (Charon *et al.*, 1997).

Orthologues of ENOD40 have also been found in other non-legume plants, suggesting roles outside of symbiosis (Compaan *et al.,* 2003). When soybean ENOD40 was expressed in wheat germ extracts, two peptides with sizes of 12- (peptide A) and 24 (peptide B) amino acid residues were formed (Röhrig *et al.,* 2002). These two peptides interacted with sucrose synthase, and therefore may be regulators of sucrose metabolism in plants (Thummler & Verma, 1987).

1.3.1.10 Nodule-specific Cysteine-rich

Nodule-specific cysteine-rich (NCR) is a gene family with more than 300 members restricted to galegoid legumes, including *M. Truncatula,* pea (*Pisum sativum*), broadbean (*Vicia faba*), white clover (*Trifolium repens*) and *Galega orientalis* (Mergaert *et al.,* 2003). The expression of gene family members has been limited to nodules and roots, with the exception of two genes that are expressed in mycorrhizal roots. NCRs encode small peptides (60-90 amino acids) with a conserved signal peptide and a conserved cysteine motif. Mergaert *et al.,* (2003) have proposed that NCRs are involved in nodule development and functioning on the basis of the abundance of transcripts in nodules and the localization of NCRs to nodules. Furthermore, they have suggested that NCRs may have a role in plant cell-to-cell signalling and plant-bacteria interactions.

1.3.2 Peptide Phytohormones with Defense Signalling Roles

1.3.2.1 Systemin

Systemin (Pearce *et al*., 1991) was the first plant peptide shown to have a role in plant signalling. This peptide is a systemically-acting, phloem-mobile wound signal capable of inducing proteinase inhibitors in unwounded tissue at femtomolar concentrations (Ryan, 1992; Schaller & Ryan, 1995). The signal is an 18-amino acid polypeptide (AVQSKPPSKRDPPKMQTD) (Pearce *et al*., 1991). Systemin is the proteolytic COOH-terminal product of the 200-residue prosystemin

precursor (McGurl *et al*., 1992), which is synthesized in the cytosol and compartmentalized in the nucleus of phloem parenchyma cells (Narváez-Vásquez & Ryan, 2004). Recently, a two-step mechanism of activation of a putative receptor was revealed. Firstly, the N-terminus of the polypeptide binds and secondly, an activation step occurs linking it to the C-terminus (Meindl *et al*., 1998). Systemin-activated signalling cascades are lipid based and are initiated by the release of linoleic acid from membranes feeding into the octadecanoid pathway. One of the components of the pathway, the oxylipins, are structural orthologues of prostaglandin and the similarities between the defense signalling pathways in tomato leaves and macrophages and mast cells led to the speculation of common ancestry between plant and animal defense pathways (Bergey *et al*., 1996). Systemin is primarily a wound response hormone, may have several other functions, including that of general defense against pathogens. In tomato leaves, exogenous systemin induced accumulation of H_2O_2, a defense product against herbivores and pathogens (Orozco-Cardenas & Ryan, 1999).

1.3.2.2 Cell Wall Hydroxyproline-rich Glycopeptides (Systemin-like Peptides)

Similar to systemin, hydroxyproline-rich glycopeptides (HypSys; Narváez-Vásquez *et al.,* 2005) are the proteolytic products of a precursor protein called preproHypSys, with 146, 165 and 214 amino acid residues in tomato, tobacco and petunia, respectively (Pearce *et al.*, 2007; Pearce & Ryan, 2003). In petunia, the 214-amino-acid precursor protein contains a leader sequence and four predicted proline-rich peptides of 18-21 amino acids, three of which bear hydroxylated prolines and glycosyl residues (Pearce *et al.,* 2007). The peptides contain a –GR- motif at their NH_2-termini, suggesting a processing site for peptide maturation (Pearce *et al.,* 2007). Wounding induces the production of the precursor in phloem parenchyma cells, which later is secreted to the cell wall matrix, where it is subjected to processing proteases (Narváez-Vásquez *et al.,* 2005).

Some of the proline residues are converted to 4-hydroxyproline via hydroxyprolination to produce mature HypSys (Matsubayashi & Sakagami, 2006).

Biochemical studies have shown that there are three HypSys glycopeptides in tomato (TomHypSys I, II and III with 20, 18 and 15 amino acid residues, respectively) and two in tobacco (TobHypSys I and II, both containing 18 amino acid residues) (Pearce & Ryan, 2003). Pearce & Ryan (2003) have demonstrated that tomato preproHypSys is upregulated on wounding and with systemin and methyl jasmonate treatment, and that TomHypSys (I, II and III) peptides, similar to systemin, act as a signal for the induction of proteinase inhibitors in leaves. Therefore, it has been suggested that systemin and HypSys peptides are wound response signals in plants (Narváez-Vásquez *et al.*, 2005).

1.3.2.3 Rapid Alkalinization Factor (RALF)

Another peptide with 49 amino acid residues was identified during systemin purification, which causes a rapid alkalinization of the medium, and was thus named rapid alkalinization factor (RALF; Pearce *et al.*, 2001). A tobacco cDNA encoding a preproprotein of 115 amino acid residues with identity to RALF at its COOH-terminus was isolated, suggesting that RALF is a processed peptide with an NH_2-terminal signal peptide (Pearce *et al.*, 2001). Sequence comparison of RALF precursors indicated the presence of an acidic region between the signal peptide and the mature RALF peptide, which may be involved in signalling (Olsen *et al.*, 2002). The presence of a dibasic (-RR-) amino acid residue within the RALF acidic region led Pearce *et al.*, (2001) to propose this site as a protease recognition site, similar to that proposed for PSK (above). RALF sequence possess a highly conserved 17-amino-acid COOH-terminal sequence motif containing two cysteine residues that have been found in many plant species, suggesting an essential role (Pearce *et al.*, 2001; Olsen *et al.*, 2002). Although Arabidopsis has 34 genes

encoding RALF-like peptides (Olsen et al., 2002), their physiological functions remain elusive (Matsubayashi & Sakagami, 2006).

1.3.2.4 AtPep1

AtPep1 (ATKVKAKQRGKEKVSSGRPGQHN) is a 23-amino-acid peptide derived from a 92-amino-acid precursor protein called PROPEP1 found in *Arabidopsis* (Huffaker *et al.,* 2006). AtPep1, similar to RALFs, can be detected via the rapid alkalinization assay. The peptide acts as an elicitor, leading to the activation of the defensin genes (PDF1.2, PR-1) subsequent to the synthesis of H_2O_2, turning on the innate immune response in plants (Huffaker *et al.,* 2006; Huffaker & Ryan, 2007).

The gene that encodes the precursor can be induced by wounding, methyl jasmonate and ethylene, suggesting a role in plant-pathogen interactions (Huffaker *et al.,* 2006; Huffaker & Ryan, 2007; Ryan *et al.,* 2007). Huffaker *et al.,* (2006) have further demonstrated that the constitutive expression of PROPEP1 results in the constitutive transcription of PDF1.2 and plant root resistance to a root pathogen, *Phythium irregulare*. Homologues of PROPEP1 have been found in Arabidopsis (six annotated and one unannotated paralogue) and in other species, making them an interesting target for the development of pathogen-resistant crops. Transcript analysis has demonstrated that PROPEP2 and PROPEP3 are the most responsive to pathogens and salicylic acid (Huffaker & Ryan, 2007).

1.3.3 Plant Natriuretic Peptides

Since systemin, was first described in 1991 (Pearce *et al*., 1991), a small number of other protein/peptide molecules have been shown to have key roles in various aspects of plant growth and development (Lindsey *et al*., 2002). A novel class of plant proteins, called plant natriuretic

peptides (PNP), with immunological similarity to vertebrate natriuretic peptides was identified by Gehring (Gehring 1999; Gehring and Irving 2003).

1.4 Natriuretic Peptide System in Vertebrates

Natriuretic peptides (NP) were first identified in vertebrate species ranging from elasmobranchs and mammals where they play a role in the regulation of salt and water balance (Maryani *et al*., 2001). Their levels are generally increased in the setting of volume expansion and act on multiple effector systems to cause vasodilation and natriuresis in an effort to return volume status back to normal (Joffy & Rosner, 2005). This regulation is partly mediated by intracellular changes in cyclic GMP (cGMP) (Gehring, 1999).

Atrial natriuretic peptide (ANP), brain natriuretic peptide (also known as B-type natriuretic peptide; BNP), C-type natriuretic peptide (CNP), ventricular natriuretic peptide (VNP) (Rosenzweig & Seidman, 1991; Farrell & Olson, 2000), D-type natriuretic peptide, and their prohormones, along with urodilatin (also known as renal natriuretic peptide), compose the major members of the natriuretic peptide family in vertebrates (shown in Table 1.1 below). The classification is based on the genetic and structural homology of their amino acid composition (Rosenweig & Seidman, 1991). Adrenomedullin is a separate peptide that also has a role in volume homeostasis. A host of other peptides, such as guanylin, uroguanylin, melanocyte stimulating hormone, and a putative endogenous ouabain-like factor also likely have a role in the regulation of body volume (Humphreys & Valentin, 2000).

ANP, BNP and VNP are cardiac hormones secreted into the blood in response to volemic, osmotic and pressure load to the heart and act to decrease the load while CNP is principally a paracrine factor in the brain and periphery (Itoh & Nakao, 1994; Fowkes & McArdle, 2000).

Table 1.1: Types of Natriuretic Peptides, their Site of Production, Stimulus and Primary Effect.

	Site of Production	***Stimulus***	***Primary Effect***
ANP	Cardiac atria	Atrial wall tension and stretch	Reduction in plasma volume and blood pressure
BNP	Cardiac ventricle	Ventricular wall tension	Same as ANP
CNP	Heart, brain, kidney, vasculature	Shear stress	Vasodilation, central nervous system neurotransmitter
D-Type Natriuretic Peptide	Unknown	Unknown	Vasodilation
Adrenomedullin	Adrenal medulla, cardiac ventricles, lungs, kidneys	Unknown	Reduction in plasma volume, blood pressure, vasodilation

Adapted from Joffy & Rosner (2005)

1.4.1 Structure

Natriuretic peptides are produced as pre-prohormones, which are cleaved to their mature form upon release (Takei, 2000) by either the enzyme furin or corin. Furin cleaves pro- CNP (Wu *et al*., 2003) and corin cleaves pro-ANP and pro-BNP (Wu & Wu, 2003). Vertebrate ANPs are synthesised from a single precursor called preproANP which consists of 149 to 153 amino acids

depending on the species. The 24 amino acid signal peptide at the N-terminal and the two carboxyl (C)-terminal arginine residues found in some vertebrates (except humans), are cleaved off to yield Pro-ANP (1 - 126). The subsequent proteolytic steps yield a single, inactive 98-amino acid called the NH_2 – terminal fragment or ANP (1 - 98) and a bioactive 28-amino acid ANP, called the COOH – terminal peptide ANP (99 - 126) (Schwartz *et al.,* 1985; Michener *et al*., 1986).

The C-terminal 28-mer, ANP (99 -126) forms a 17-residue circular structure due to disulphide bonds formed between the cysteines in positions 7 and 23 and is believed to be the main biologically active ANP (Inagami *et al.,* 1987). This 17-amino acid ring structure shares a high degree of sequence homology between members of the natriuretic family. It has been demonstrated that ANP, BNP and CNP all contain this loop, but the amino acid sequence of this loop is identical in ANP and BNP, but CNP differs in residues 3 and 5 (Koller *et al*., 1991; Sudoh *et al*., 1988; Sudoh *et al*., 1990). ANP (1 -98) is less well established, but it has been suggested that it is further cleaved into ANP (1 - 30), ANP (31 - 67) and ANP (79 - 98), with a role in Na^+ and K^+ excretion and long acting effects on vascular smooth muscles (Vesely *et al*., 1994). The four peptide hormones consisting of amino acids 1 – 30, 31 – 67, 79 – 98 and 99 – 126 of the ANP prohormone have been tentatively named long-acting natriuretic peptide, vessel dilator, kaliuretic peptide, and atrial natriuretic peptide, respectively, for their most prominent effects in humans (Vesely *et al.,* 1994; Vesely *et al.,* 1998) and laboratory rats (Martin *et al.,* 1990) as shown in figure 1.1 below.

Structural divergence between the natriuretic peptides is largely restricted to differences in the N- and C-terminal portions and probably has an important role in differentiating the function of the

particular peptide. For example, urodilatin (renal natriuretic peptide) differs from ANP only in the addition of 4 amino acids in the N-terminal region (Feller et al., 1989).

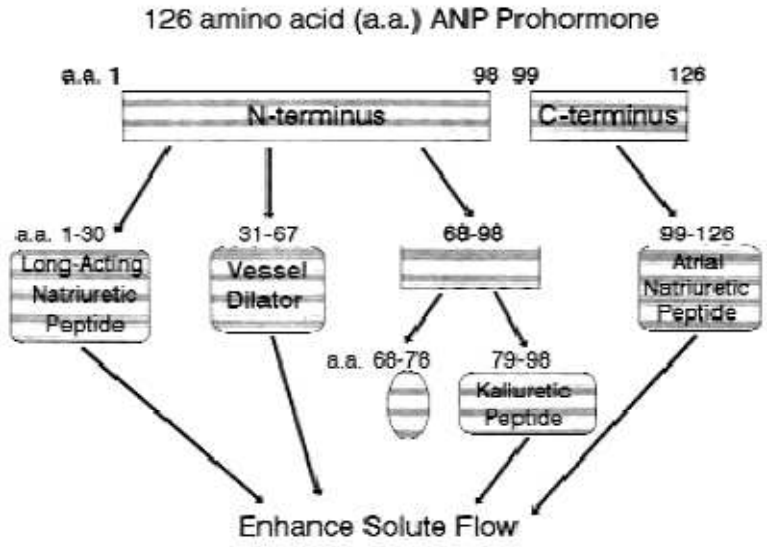


Figure 1.1. Origin of long-acting natriuretic peptide, vessel dilator, kaliuretic peptide and atrial natriuretic peptide from the 126 amino acid atrial natriuretic peptide (Vesely *et al.*, 2001)

1.4.2 Natriuretic Peptide Receptors (NPR-A, NPR-B, NPR-C & NPR-D)

The effects of NPs are mediated through interaction with receptors (NPRs) and four types of receptors are involved as shown in table 1.2 below. The four types of natriuretic peptide receptors include the following;

- NPR-A and NPR-B
- NPR-C
- NPR-D

NPR-A and –B contain an intracellular guanylate cyclase domain (Garbers & Lowe, 1994), of which binding of the ligand to the extracellular domain stimulates intracellular production of cGMP (Takei, 2000). NPR-A exhibits high affinity to ANP and BNP and NPR-B is specific to

CNP (Rosenzweig & Seidman, 1991; Takei, 2000). It is interesting to note that VNP has high affinity to both NPR-A and NPR-B (Katafuchi *et al*., 1994). NPR-C lacks a cytosolic GC domain and thus does not use cGMP as a second messenger to exert biological actions. However, NPR-C is believed to be linked to cAMP pathways and also functions as a 'clearance' receptor by internalizing and metabolizing NPs (Anand-Srivastava & Trachte, 1993). The supportive data for its clearance function are 1) no ligand specificity (high affinity to all NPs); 2) overall tissue distribution (more than 90% of total NP receptor population in various tissues); and 3) rapid internalization after binding of the ligand to the receptor. The NPR-C is particularly abundant in the lung of tetrapods and in the gills of teleost fishes. These organs are ideally located to serve as a buffer to regulate the plasma concentration of hormones secreted from the heart, since they receive the entire blood expelled from the heart (Olson, 1998).

A new type of GC-uncoupled receptor has been cloned from the brain of eels in addition to NPR-A and NPR-C from the gills and NPR-B from the brain, and named NPR-D (Kashiwagi *et al*., 1995). The GC-coupled receptors, NPR-A and NPR-B, exist as tetramers under non-reducing conditions while NPR-C is a dimeric receptor. However, NPR-D is tetrameric similar to GC-coupled receptors. HS-124-1, which was initially thought to be a specific blocker for GC-coupled receptors (Morishita *et al*., 1991), also blocked the binding of NPs to NPR-D. Therefore, it is evident that HS-124-1 is a specific blocker for tetrameric NP receptors. Since NPR-D is localized selectively in the brain of eels, it may have some biological functions more than the mere clearance of NPs, as suggested for NPR-C (Levin, 1993).

Table 1.2: Types of Natriuretic Peptide Receptors, their Second Messenger and Tissue Distribution.

Receptor	***Second Messenger***	***Tissue Distribution***
NPR-A	cGMP	Kidney, adrenal, endocardium, brain, lung, aorta
NPR-B	cGMP	Kidney, adrenal, cerebellum, pituitary, lung
NPR-C	None known	Widely distributed

Adapted from Joffy & Rosner (2005)

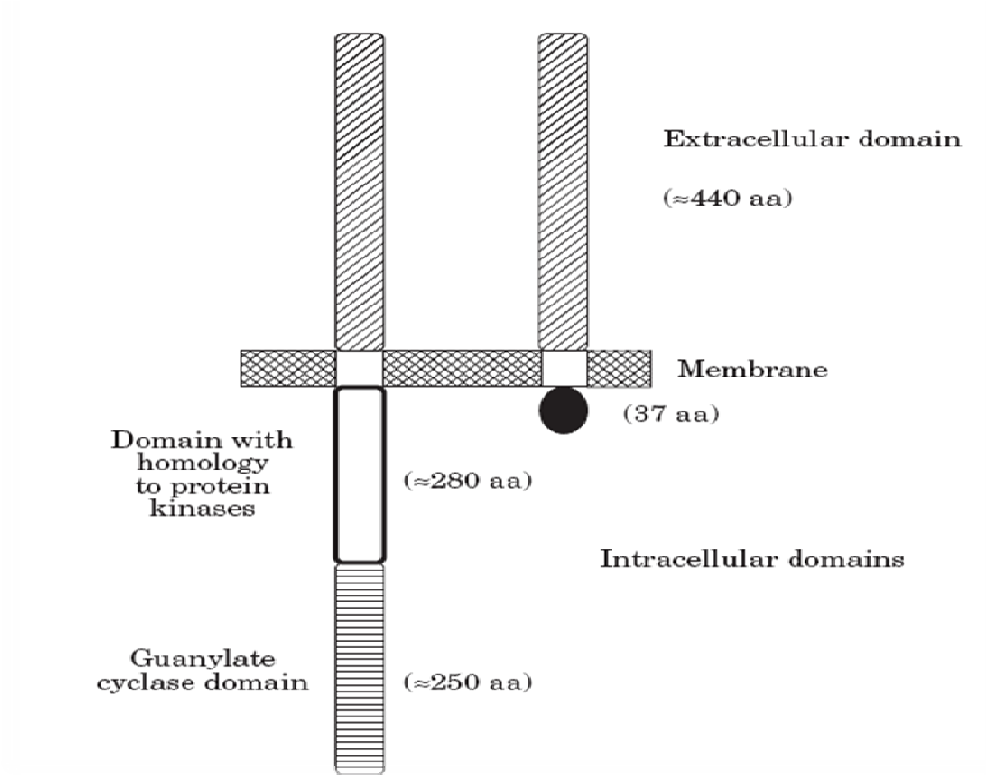


Figure 1.2: Representation of the principal two types of NP receptors, NPR-A and -B and NPR-C (Gehring, 1999).

The receptors are single transmembrane proteins with extracellular domains of approximately 440 amino acids as shown in figure 1.2 above. The intracellular domain of NPR-C is short, 37

amino acids. NPR-A and NPR-B, however, both have bipartite intracellular domains of approximately 280 and 250 amino acids respectively. The domains proximal to the membrane, which are approximately 280 amino acids, show significant homology to protein kinases while the distal fragments include guanylate cyclase domains (approximately 250 amino acids) (Takayanagi *et al.,* 1987; Schulz *et al.,* 1989; Koller *et al.,* 1991).

1.4.3 Functions and Modes of Action

Sustaining water and solute homeostasis is a key requirement for living systems and in vertebrates homeostasis is, in part achieved by natriuretic peptides (NPs), a family of peptide hormones (Suzuki *et al*., 2001). The molecular structures and physiological functions of natriuretic peptides (NPs) in animals are the subject of a large and growing literature (Kourie & Rive, 1999). This wealth of information is, in part due to the significant therapeutic potential of the peptide(s) and its derivatives.

In vertebrates, natriuretic peptides constitute a family of peptide hormones which occur not just in the heart but also the brain and periphery of vertebrates and are strongly implicated in the regulation of salt and water homeostasis. NPs also modulate cation movement by inhibiting the apical Na^+, K^+ - ATPases; by stimulating (Petrov *et al.,* 1994) or inhibiting (Carmelo *et al*., 1994) antiporters in different cell types. NPs have also been reported to also promote K^+ conductance in rat glomerular mesangial cells (Cermark *et al*., 1996). They also facilitate a K^+ current in atrial ventricular papillary muscle (Kecskemeti *et al*., 1996).

The synthesis and release of natriuretic peptides are stimulated by various derangements in systemic blood pressure, as well as increases in extracellular volume and sodium balance. They exert a multitude of endocrine and paracrine actions in the heart, brain, kidneys, vasculature and

adrenal glands. However, their major role is to induce natriuresis through their actions on renal hemodynamics and tubular function. Other effects of natriuretic peptides include vasodilation, decrease in sympathetic outflow, inhibition of arginine vasopressin release, and inhibition of aldosterone release (Humphreys & Valentin, 2000).

ANP is synthesized in the cardiac atria of vertebrates and regulates blood pressure and extravascular flow of fluids (De Bold, 1985; Lang *et al*., 1985). Furthermore, Ca^{2+}-dependant K^{+} channels in mesangial cells are activated by ANP and its putative second messenger cGMP (Stockand & Sansom, 1996). It is released in response to atria stress caused by hypervolemia – induced increased blood pressure (Lang *et al*., 1985). It is thus concluded that the heart (atria) contains the highest concentration of ANP than the rest of the body (Needleman *et al*., 1989).

BNP is synthesized and distributed in the atria and ventricle of the heart, even though it was first isolated from brain tissue (Sudoh *et al*., 1988). BNP is found in higher concentrations in the heart ventricle, and is considered a cardiac natriuretic peptide. BNP, like ANP has a role in vaso-relaxing, natriuretic and diuretic responses (Pidgeon *et al*., 1996; Rademaker *et al*., 1997).

CNP is found in a wide array of vertebrate tissue (de Plater *et al*., 1998). It was thought to only occur in the central nervous system, maintaining body fluid homeostasis (Sudoh *et al*., 1990), but has been discovered in rat heart (Vollmar *et al*., 1993) and as a native peptide in hearts and brains of elasmobranch fishes (Suzuki *et al*., 1992). CNP contributes to the regulation of vascular tone, growth and hormone release (Shimekake *et al*., 1994; Amin *et al*., 1996).

A unique peptide named VNP has been isolated from the cardiac ventricle of two teleost species. There is a possibility that VNP is a homologue of tetrapod BNP, but it is structurally more akin to

mammalian ANP than to BNP (Takei *et al.,* 1994). VNP has a long C-terminal 'tail' sequence with 14 amino acid residues that extend from the intramolecular ring (Takei, 2001).

There is structural and functional evidence of an immunologically related peptide that operates in plants (Vesely *et al*., 1993; Billington *et al*., 1997; Gehring, 1999). Natriuretic peptide immunoanalogues occur in many plants and have been isolated, with two NP encoding genes identified in *Arabidopsis thaliana* L. *(AtPNP-A* and *AtPNP-B)* (Ludidi *et al*., 2002).

1.5 Biologically Active Natriuretic Peptide System in Plants

Plant natriuretic peptides represent a novel class of small proteins that are most closely related to expansins, which are key regulators of cell wall extension (Ludidi *et al*., 2002; Li *et al*., 2003; Kende *et al*., 2004).

1.5.1 Evidence for a Natriuretic Peptide System in Plants

Immunological and physiological evidence suggests the presence of biologically active natriuretic peptide hormones (NPs) in plants (Gehring & Irving, 2003).

1.5.1.1 ^{125}I – labeled rANP Binds to Isolated Leaf Microsomes

The first indications for NPs in plants came from radioimmunoassays on Florida beauty (*Dracena godseffiana*) (Vesely & Giordano, 1991) where antibodies against the N-terminus (ANP, 1–98), the mid-portion (ANP, 31–67) and the C-terminus (ANP, 99–126) recognized peptides in leaves and stems. It was also demonstrated that a synthetic peptide identical to the C-terminus (amino acids 99-126) of the rat atrial natriuretic peptide (rANP) binds specifically to isolated leaf microsomes *in vitro* (Gehring *et al.,* 1996) and leaf tissue *in situ* (Suwastika *et al.,* 2000).

Evidence for receptor ligand interactions comes from binding studies where isolated leaf membranes were exposed to rat (3-[^{125}I] iodotyrosol^{28}) ANP (Gehring *et al*., 1996; Suwastika *et al*., 2000). The specificity of binding was assessed by increasing the concentration of unlabelled competitor (rANP). Approximately 50% of the labeled ligand was displaced by 0.1 µM rANP, indicating that plant membranes contain a low-affinity NP binding site for the heterologous rANP.

The rate of transpiration, solute flow and solute uptake in carnation and chrysanthemum was shown to rapidly and significantly increase after exogenous application of synthetic human ANP (1 - 30), ANP (31 - 67) or ANP (79 - 98), at concentrations of less than 5.9 pg.ml^{-1} (Vesely *et al*., 1993). However, ANP (99 - 126) was shown not to affect these processes at equivalent peptide concentrations.

1.5.1.2 rANP Promotes Stomatal Opening

Secondly, it was also demonstrated that synthetic rANP can induce stomatal opening in *Tradescantia* sp. at concentrations (of $\geq 10^{-6}$ M) in a conformation dependent manner (Gehring *et al*., 1996; Pharmawati *et al.*, 1998a and 2001). The high concentration necessary to get a response could be due to large taxonomic divergence between hormone source and test organism. It is noteworthy that Na^+ is required in the medium for activity in animal systems, which is not the case in plants (Gehring *et al*., 1996). This suggests that in plants NPs operate on processes other than Na^+ transport, such as K^+ transport or the synthesis of compatible solutes.

1.5.1.3 cGMP

Thirdly, there are indications that this NP effect on stomatal guard cells is influenced by cGMP since stomatal opening does not occur in the presence of LY 83583, an inhibitor of particulate guanylate cyclase, but it can be induced by the cell permeant cGMP analogue 8-Br-cGMP (Pharmawati *et al.*, 1998a, b and 2001). For decades the presence and possible function of cyclic nucleotides in plants was a subject of fierce debate (Newton *et al.*, 1999) however, Ludidi and Gehring (2003), described the first plant guanylate cyclase and more recently new evidence for the importance of cGMP to plant natriuretic peptide signalling was reported (Wang *et al*., 2007).

Three different guanylate cyclase inhibitors that have different modes of action were used namely; LY83583, ODQ and NS 2028. LY 83583 is a redox-sensitive dye that is a cell permeable competitive inhibitor of soluble guanylate cyclases that acts by generating superoxide anion radicals (Mülsch *et al*., 1988). ODQ and NS 2028 are analogues that are specific irreversible inhibitors of soluble guanylate cyclase *in vitro* thought to act at the haem-binding pocket and hence can be reversed in intact cells (Garthwaite *et al*., 1995; Schrammel *et al*., 1996; Olesen *et al*., 1998). Of these inhibitors, LY83583 has been used most often in plant systems (Penson *et al*., 1996; Pharmawati *et al*., 2001). It is critical to use several inhibitors to assure that heterologous pharmacological agents do not cause misleading results. Because two of these inhibitors reduced *At*PNP-A and all three inhibited the peptide region 33 - 66 induced increases in protoplast volume, it strongly supports a role for guanylate cyclase activity in relaying the PNP signal. The membrane permeable analogue of cGMP, 8-Br-cGMP, was shown to increase protoplast volume at 0.1 μM but that it had no effect at 10 μM. This could be due to an overriding feedback system at the higher concentrations of 8-Br-cGMP or possibly toxicity (Wang *et al*., 2007).

It was also shown that, rANP increased radial water movement out of the xylem shoots of *Tradescantia multiflora* (Suwastika *et al*., 2000), by measuring the dye movements and determining the water tissue ratios by ^{2}H NMR. An increase in radial movements was observed when an increase in 8-Br-cGMP occurred, while the channel inhibitor mercuric chloride and LY 83583 (guanylate cyclase inhibitor) both significantly inhibited the radial movement. This evidence suggests that NPs play an important role in controlling radial movements out of the xylem and is mediated by the regulation of guanylate cyclase and water channels (Gehring, 1999).

1.5.1.4 Isolation & Purification of Active PNP immuno-analogues

Fourthly, and most importantly, biologically active plant natriuretic peptide immuno-analogues (irPNP) have been isolated and purified by immunoaffinity chromatography. IrPNPs from ivy (Hedera helix) were the first such NPs from a plant, using rabbit anti-alpha-ANP (1 - 28) (human, canine) antiserum (Billington *et al.*, 1997). Isolation and affinity purification of immunoreactive plant NPs (irPNPs) from ivy (*Hedera helix*) and potato (*Solanum tuberosum*) with rabbit anti-α-ANP (99–126) (human) antiserum yielded several different molecules, indicating that plants either contain more than one type of NP or that different precursors also contain the epitope(s) (Pharmawati *et al*., 1998a).

Two *Arabidopsis thaliana* irPNP-encoding genes termed *AtPNP-A* and *AtPNP-B* have been identified and the domain organization of the encoded proteins described (Ludidi *et al.*, 2002). Both proteins are novel and have not been functionally characterized fully. However, *AtPNP-B* has an orthologue in *Citrus jambhiri* that is associated with responses to citrus blight (Ceccardi *et al*., 1998).

From the evidence above, it is noted that, irPNPs promote stomatal opening (Pharmawati *et al*., 1998a), affect ion transport in plants (Pharmawati *et al*., 1999), displace ANP-binding (*in situ* and *in vitro*) (Suwastika *et al*., 2000) and cause rapid, reversible cGMP increases (Pharmawati *et al*., 1998b; Pharmawati *et al*., 2001). The responses to irPNP and ANP showed a high degree of similarity.

1.5.2 Structure

Recently, progress has been made in the elucidation of the molecular structure of irPNP-like molecules. In *Arabidopsis thaliana* two irPNP encoding genes (*AtPNP-A* and *AtPNP-B*) have been identified and characterized (Ludidi *et al.*, 2002). AtPNP-A and –B are small proteins of less than 130 amino acids and are related to expansins. AtPNP-A is a small protein of 126 amino acids in length (MW: 14016 kDa; pI: 9.22) that is encoded by a gene with a single intron of 100 bp (accession no.: AAD08935), as shown in figure 1.3 below. The DNA sequence of the AtPNP-A coding region is also shown in figure 1.4 below. The protein contains a 24 amino acid signal peptide (MW: 2249).

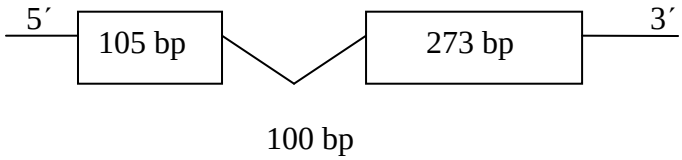


Figure 1.3. Intron and exon organization of the A*tPNP-A* encoding gene from *Arabidopsis thaliana*

1 ATG GCG GTG AAA TTT GTG GTG GTG ATG ATT GTG TTT GCG CAG ATT 45
46 CTG GCG CCG ATT GCG GAA GCG GCG CAG GGC AAA GCG GTG TAT TAT 90
91 GAT CCG CCG TAT ACC CGC AGC GCG TGC TAT GGC ACC CAG CGC GAA 135
136 ACC CTG GTG GTG GGC GTG AAA AAC AAC CTG TGG CAG AAC GGC CGC 180
181 GCG TGC GGC CGC CGC TAT CGC GTG CGC TGC ATT GGC GCG ACC TAT 225
226 AAC TTT GAT CGC GCG TGC ACC GGC CGC ACC GTG GAT GTG AAA GTG 270
271 GTG GAT TTT TGC CGC GAA CCG TGC AAC GGC GAT CTG AAC CTG AGC 315
316 CGC GAT GCG TTT CGC GTG ATT GCG AAC ACC GAT GCG GGC AAC ATT 360
361 CGC GTG GTG TAT ACC CCG 378

Figure 1.4. Nucleotide sequence of AtPNP-A from *Arabidopsis thaliana* (378 base pairs).

1.5.3 Synthesis and Degradation

The loci of synthesis and degradation as well as induction patterns of irPNPs are currently under investigation. Reverse transcriptase-PCR (RT-PCR) has revealed *AtPNP- A* transcripts in leaf tissue from unstressed *Arabidopsis* plants (Ludidi *et al*., 2002), indicating that the protein is not only found, but synthesized, in leaves. It is also possible that stress induced up-regulation of *AtPNP-A* transcription occurs either in leaves or elsewhere. Tissue prints revealed that immunoreactants are concentrated in vascular tissues of leaves, petioles and stems. Phloem-associated cells, xylem cells and parenchymatic xylem cells showed the strongest immunoreaction (Maryani *et al.*, 2003).

In situ localization data indicates that irPNPs are systemically mobile proteins since they are found in conductive tissue; a highly unlikely place for synthesis, but not transport. Biologically active immuno-reactants have also been detected in xylem sap and this is indicative of synthesis in the root. Further information on temporal and spatial induction patterns will be forthcoming when transgenic plants containing GUS or GFP promoter fusions are available (Maryani *et al.*, 2003).

1.5.4 Functions of Plant Natriuretic Peptide (PNP)

There is evidence that irPNPs have a role in plant homeostasis. It has been shown that levels of native irPNP increased in salt stressed *Erucastrum strigosum* (a fast growing brassicaceous weed) shoots and that irPNP levels also rise in *Arabidopsis* suspension cells exposed to NaCl or sorbitol levels (Rafudeen *et al*., 2003). This is suggestive of a role for PNP in counteracting the salinity or drought stress.

Net water up-take into plant cells is osmotic in nature and occurs in response to highly regulated ion movements. Under unfavorable osmotic conditions, for example high salinity, plants can synthesize intracellular compatible solutes such as proline or mannitol, thus increasing intracellular osmotic pressure and drawing in water. Enhanced water uptake does occur in response to NPs in plants (Gehring, 1999; Maryani *et al*., 2001) and may be the driving force of the observed stomatal opening which is conditional on water uptake. Since irPNP does not appear to change water permeability and net ion flux appears with a delay (>20 min) (Pharmawati *et al*., 1999), it is interesting to speculate that irPNPs stimulate rapid synthesis (or re-compartmentalization) of compatible solutes. Such a mode of action is in accordance with previously observed NP-dependent increases in lateral water movement out of the conductive tissue (xylem) into the neighbouring parenchyma (Suwastika & Gehring, 1998; Gehring 1999), increases in cell volumes (Maryani *et al*., 2001), the delayed net K^+ and Na^+ uptake (Pharmawati *et al*., 1999) and the promotion of stomatal opening (Gehring *et al*., 1996; Pharmawati *et al*., 1998a; Gehring, 1999). The ‘drawing’ of water into cells and tissues together with an up-regulation under conditions of drought and salinity stress are also compatible with a role for PNP in plant homeostasis.

1.5.5 Evolutionary Development

In mammals, the natriuretic peptide (NP) system was originally discovered in an extract of rat atria in 1981 (de Bold *et al*., 1981). As mentioned previously, there is some evidence that the natriuretic peptide system is also present in plants. The evidence of a natriuretic peptide system was shown from a wide range of diversified plant species. *At*PNP-A displays a weak overall sequence identity with human ANP, although significant similarity is found between residues 33 – 66 of *At*PNP-A and the mature human ANP. This section of the peptide (33 - 66) is also sufficient to confer solute uptake by the protoplast (Morse *et al*., 2004). Plant natriuretic peptide genes also share weak sequence similarities with the expansins, although they lack a carbohydrate-binding domain. Natriuretic peptides represent the first known conserved hormonal system that is shared between animals and plants. The sequence similarity and homology in the signalling mechanism strongly support that this mechanism, involved in homeostasis regulation, may have evolved before the last common ancestor of the two kingdoms, 1.6 billion years ago (Meyerowitz, 1999).

Humans and plants share a convergent evolution rather than a common inheritance which is suggested by the fact that green plants took a radically different evolutionary path in which movement was a low priority, but, however, there are traits that are shared by humans and plants such as multi-cellularity and tissue differentiation. Atrial natriuretic peptides have been identified in both metazoans and green plants, and serve to regulate solute flow through tissues in both groups (Billington et al., 1997).

In silico analysis has also established the evolutionary and functional relationships of irPNP-like molecules within the superfamily of expansins, pollen allergens and distantly related molecules such as endoglucanases and shown that irPNP-like molecules are related to expansins (Ludidi *et*

al., 2002). Gehring & Irving (2003) also hypothesised that irPNP-like molecules have evolved from primitive glucanase-like molecules that have been recruited to become systemically mobile modulators of homeostasis acting via the plasma membrane. Such a function is compatible with localisation in the conductive tissue and the physiological and cellular modes of action of irPNPs (Gehring & Irving, 2003).

1.6 Molecules Related to Plant Natriuretic Peptides

Plant natriuretic peptides represent a novel class of small proteins that are most closely related to expansins, which are key regulators of cell wall extension (Ludidi *et al*., 2002; Li *et al*., 2003; Kende *et al*., 2004). A significant sequence similarity of *At*PNP-A (AAD08935) and *At*PNP-B (CAB79756) to a functionally uncharacterised transcript from citrus, CjBAp12 (AAD03398) has been observed (Ludidi *et al*., 2002). AtPNP-A and AtPNP-B, as well as CjBAp12, show homology with the *N*-terminus of expansins but are considerably shorter lacking the wall binding domain (Ludidi *et al*., 2002; Li *et al*., 2003). However, unlike expansins, PNPs are mobile and seem to be involved in regulating solute and water homeostasis (Gehring & Irving 2003; Maryani *et al*., 2003; Rafudeen *et al*., 2003; Morse *et al*., 2004).

1.6.1 Expansins

Plant cells are surrounded by a rigid cell wall. The major components of the wall are cellulose microfibrils consisting of an unbranched β-1,4-glucan polymer, hemicelluloses (branced glycans) that bind to cellulose to form a matrix, and acidic polysaccharides (pectins) that form ionic gels (Ludidi *et al*., 2002). Developmental processes such as cell elongation and any rapid elongation such as stomatal movement that require changes in cell shape thus necessitate temporary loosening of the cell walls.

Cell wall loosening and extension play a crucial role in shaping the form and size of plants and a highly conserved family of proteins called expansins, that catalyzes wall extension have been identified. Expansins are cell wall proteins that induce pH-dependent wall extension and stress relaxation in a characteristic and unique manner. Expansins were first discovered in studies of the mechanism of plant cell enlargement, specifically of the "acid growth" phenomenon (McQueen-Mason *et al.*, 1992). Evidence from work with transgenic plants indicates that endogenous expansins are not only involved in regulating growth (Cho & Cosgrove, 2000; Choi *et al.*, 2003; Zenoni *et al.*, 2004), but also developmental processes such as initiation of leaves (Pien *et al.*, 2001), abscission (Cho & Cosgrove, 2000), and softening of tomato fruits (Brummell *et al.*, 1999).

Two classes of expansin proteins are currently recognized, namely alpha-expansins (EXP) and beta-expansins (EXPB), and they comprise large multigene families whose members show diverse organ-, tissue- and cell-specific expression patterns. They are similar to each other in size (~ 25 – 28 kDa for the mature protein) and have distant but significant sequence similarity to each other throughout the length of the protein backbone (Cosgrove *et al*., 2002).

A number of sequence features are common to the two expansin families, including a cleaved signal peptide at the amino terminus, a series of cysteines with characteristic spacing and conserved flanking sequences, an "HFD" (histidine, phenyalanine, glutamic acid) motif, and a series of tryptophans and other aromatic residues at characteristic positions in the protein backbone. The two families also differ in the presence of N-linked glycosylation motifs, which are generally absent in the EXP sequences but present in EXPB sequences (Cosgrove *et al*., 2002).

Sequence analysis suggests that mature expansins contain two domains. Domain 1 has significant but distant homology to family-45 endoglucanases, including a series of conserved cysteines and an HFD motif that makes up part of the catalytic site of family-45 endoglucanases (Cosgrove, 1997; Cosgrove, 2000). This suggests that expansins have an endoglucanase activity, but if so it is very weak and highly specific in its substrate requirements because experimental tests of EXP have not detected significant activity of its kind (McQueen-Mason & Cosgrove, 1995). It would thus appear that the structural similarity between glucanases and expansins reflects an evolutionary relationship between two classes of enzymes that operate differently on a common substrate, the cell wall (Ludidi *et al*., 2002). Domain 2 has homology to a group of proteins found in grass pollen, named group 1 grass pollen allergens, of unknown biological and biochemical function. The group 1 allergens are glycoproteins abundantly expressed on the surface of pollen grains, released upon hydration and have been identified as the main causative agents of hay fever and allergic asthma elicited by grass pollen (Knox & Suphioglu, 1996). It has been speculated that domain 2 is a polysaccharide binding domain on the basis of conserved aromatic and polar residues on the surface of the protein (Cosgrove, 1997).

In *Arabidopsis* the EXP family has 26 members and the EXPB family has five members (Li *et al*., 2002), whereas in rice and in maize there are probably a similar number of EXP genes but many more EXPB genes (Lee *et al*., 2001; Lee & Kende 2002; Wu *et al*., 2001). The larger size of the EXPB family in grasses is likely related to the unusual composition of the grass cell wall, which had reduced amounts of xyloglucan and pectins and larger amounts of glucuronoarabinoxylans (Carpita, 1996). The evolutionary divergence of the composition of the grass cell wall appears to have been accompanied by an enlargement and functional diversification of the EXPB family (Cosgrove *et al*., 2002).

In *Arabidopsis* there is also a third family of related proteins, named expansin-like (EXPL), as well as a single, more distant protein named expansin-related (EXPR) (Cosgrove *et al*., 2002). Additionally, *Arabidopsis* and other species have genes that encode homologs of a protein first found in studies of citrus blight (Ceccardi *et al*., 1998). This protein, called p12 (CjBAp12) because its size is approximately 12 kDa, has very distant homology to Domain 1 of EXP and is thus considerably smaller than the classic expansins (molecular mass of ~ 25 kDa) (Cosgrove *et al*., 2002). The protein accumulates in root, stem, and leaf tissues (Ceccardi *et al*., 1998), suggesting systemic mobility. However, despite the sequence similarity with expansins, p12 has no apparent expansin-like activity suggesting a role other than wall loosening (Ceccardi *et al*., 1998).

The plant natriuretic peptides share sequence homology with expansins but do not contain the tryptophan and tyrosine rich C-terminal putative polysaccharide-binding domain typical of expansins or bacterial cellulases and hemicellulases (Ludidi *et al.*, 2002). Interestingly, the C-terminal domain of α and β expansins (Cosgrove, 2000) is encoded by an entire exon that is absent in irPNP-like molecules (Ludidi *et al*., 2002). Expansins are distantly related to glucanases and cellulases; in the latter, the C-termini have been proven to be cell wall binding (Linder & Teeri, 1997) and the same function has been suggested for the α and β expansin C-terminus (Cosgrove, 2000). Since expansins, which are the closest relatives of irPNPs, and the more distantly related glucanases and cellulases, all contain the C-terminus; it is reasonable to argue that irPNP-like molecules have lost this domain.

It can also be argued that both irPNP-like molecules and expansins have evolved from ancestral glucanase-like molecules that hydrolyzed the cell wall. The absence of the polysaccharide-binding domain is thought to increase the mobility of these molecules (Ludidi *et al.*, 2002).

Importantly, it has previously been demonstrated that irPNPs act on protoplasts (Maryani *et al*., 2001), that is plant cells without cell walls as well as microsomes, indicating that these novel proteins specifically interact with the plasma membrane. Thus, irPNP-like molecules are candidates for a physiological role in water and solute homeostasis (Gehring & Irving 2003; Maryani *et al*., 2003; Rafudeen *et al*., 2003; Morse *et al*., 2004).

1.6.2 Citrus Blight (CjBAp12/p12)

The *AtPNP-A* and *AtPNP-B* encoded proteins, which are both related proteins, show similarity to CjBAp12, a functionally undefined protein isolated from *Citrus jambhiri* (rough lemon). CjBAP12 (previously referred to as a blight-associated protein, p12) is a protein that is induced in response to citrus blight infection and has a molecular mass of approximately 12 kDa (Ceccardi *et al*., 1998). The CjBAp12 gene is expressed in the root and stem tissues in response to a challenge from citrus blight, which proliferates in the conductive tissue of the host and severely affects host homeostasis eventually resulting in xylem plugging and consequent shoot wilting and host death. The protein itself accumulates in root, stem, and leaf tissues (Ceccardi *et al*., 1998), suggesting systemic mobility.

IrPNPs and CjBAp12 are likely to be not just structural homologues but may also have some common physiological functions. Since CjBAp12 is induced by citrus blight which leads eventually to die-back, it is conceivable that the expression of CjBAp12 is an early host response to counteract the pathogen induced limitation of water and nutrient availability (Nembaware *et al*., 2004). If so CjBAp12 may, at least in part, enable the host to counteract blight-induced disturbances of cation and water transport. It is noteworthy, that the CjBAp12 message is present in roots but not leaves while the protein is present in xylem sap and leaves (Ceccardi *et al*., 1998).

This must mean that the message in leaves has an extremely short half-life or more likely, that the protein is transported, presumably in the xylem, to the leaves.

On the basis of cross reaction with antibodies raised against vertebrate atrial natriuretic peptide, it was speculated that p12-like proteins may function as peptide hormones for water and solute balance in plant cells (Ludidi *et al*., 2002).

How then are the irPNP-like proteins and p-12, interrelated and related to other proteins, and can similarities with proteins of known function, therefore provide a key to understanding their unknown function?

1.7 Pathogenesis - Related Proteins

Pathogenesis-related proteins (PRs) are a group of plant-coded proteins induced by different stress stimuli. PRs have been assigned an important role in plant defense against pathogenic constraints and in general adaptation to stressful environment (Edreva, 2005).

Since their discovery in tobacco leaves hypersensitively reacting to tobacco mosaic virus (TMV) by two independently working groups (Van Loon & Van Kammen, 1970), pathogenesis-related proteins (initially named “b” proteins) have focused an increasing research interest in view of their possible involvement in plant resistance to pathogens. Initially, it was assumed that these proteins are commonly induced in resistant plants. However, it later turned out that b-proteins are induced not only in resistant, but also in susceptible plant – pathogen-interactions, as well as in plants, subjected to abiotic stress factors (Van Loon, 1985).

1.7.1 Classification of PRs

Originally, five main groups of PRs (PR-1 to PR-5) were characterized by both molecular and molecular-genetic techniques in tobacco. The PR groups are numbered in order of decreasing electrophoretic mobility. Each group consists of several members with similar properties (Bol *et al.,* 1990). Thereupon, in 1994 a unifying nomenclature for PRs was proposed based on their grouping into families sharing amino acid sequences, serological relationships, and/or enzymatic or biological activity. Eleven families (PR-1 to PR-11) were recognized and classified for tobacco and tomato, with the families PR-8 and PR-10 being also present in cucumber and parsley, respectively (Van Loon *et al.,* 1994). Later three novel families (PR-12, PR-13 and PR-14) were recognized in radish, Arabidopsis and barley, respectively (Van Loon & Van Strien, 1999). Germins and germin-like proteins (GLPs) have been classified as PR-15 and PR-16; PR-16 has been isolated from hot pepper during the resistance response to bacterial and viral infection (Park *et al.,* 2004). PR-17 proteins have been found in infected tobacco, wheat and barley (Christensen *et al.,* 2002).

The criteria used for the inclusion of new families into PRs are that (a) the protein must be induced by a pathogen in tissues that do not normally express it, and (b) induced expression must occur in at least two different plant-pathogen combinations, or expression in a single plant-pathogen combination must be confirmed independently in different laboratories (Van Loon & Van Strien, 1999).

Within each PR-family a type member has been defined, the nucleotide sequence of the mRNA of which may be used in the search for homologues in the same or in different plant species as shown in the Table 1.3 below.

Table 1.3: Recognized families of pathogenesis-related proteins (van Loon *et al.,* 2006)

Family	***Type Member***	***Properties***
PR-1	Tobacco PR-1a	Unknown
PR-2	Tobacco PR-2	B-1,3-glucanase
PR-3	Tobacco P, Q	Chitinase type I, II, IV, V, VI, VII
PR-4	Tobacco 'R'	Chitinase type I, II
PR-5	Tobacco S	Thaumatin-like
PR-6	Tomato Inhibitor I	Proteinase-inhibitor
PR-7	Tomato P_{69}	Endoproteinase
PR-8	Cucumber chitinase	Chitinase type III
PR-9	Tobacco "lignin-forming peroxidase"	Peroxidase
PR-10	Parsley "PR1"	Ribonuclease-like
PR-11	Tobacco "class V" chitinase	Chitinase, type I
PR-12	Radish Rs-AFP3	Defensin
PR-13	Arabidopsis TH12.1	Thionin
PR-14	Barley LTP4	Lipid-transfer protein
PR-15	Barley OxOa (germin)	Oxalate oxidase
PR-16	Barley OxOLP	Oxalate-oxidase-like
PR-17	Tobacco PRp27	Unknown

For the majority of the PR families, activities are known or can be inferred (Kombrink & Somssich, 1997). The PR-2 family consists of endo-β-1,3-glucanases, and PR-3, -4, -8 and -11 are all classified as endochitinases, which could act against fungi (Brunner *et al.,* 1998). Members of the PR-8 family possess lysozyme activity and may be directed against fungi. PR-6 is a group of proteinase inhibitors implicated in defense against insects and other herbivores, micro-organisms, and nematodes (Koiwa *et al.,* 1997; Ryan, 1990). PR-7 is an endoproteinase that is the most conspicuous PR in tomato (Jorda *et al.,* 2000) and it might aid in microbial cell wall dissolution. PR-9 is a specific type of peroxidase that could act in cell wall reinforcement by

catalyzing lignifications (Passardi *et al.,* 2004) and enhance resistance against multiple pathogens.

PRs of group 5 share significant amino acid sequence homology with the sweet tasting protein in the fruits of the tropical plant *Thaumatococcus danielli*, and have been named thaumatin-like (TL) proteins (Cornelissen *et al.,* 1987). Some members of this family have been shown to possess antifungal activity, particularly against oomycetes (Van Loon & Van Strien, 1999).

Members of the PR-1 proteins have also been associated with activity against oomycetes. Notably, the prominent PR-1 proteins are often used as markers of the enhanced defensive state conferred by pathogen-induced systemic acquired resistance (SAR), but their biological activity has remained elusive (Van Loon & Van Strien, 1999).

PR-10 shows homology to ribonucleases (Bufe *et al.,* 1996) and it has sometimes been assumed that the ribonuclease activity of PR-10-type proteins points to a role in defense against viruses (Park *et al.,* 2004).

Defensins (PR-12) (Lay & Anderson, 2005), thionins (PR-13) and the lipid transfer proteins (LTPs; PR-14) all exhibit antibacterial and antifungal activities, exerting their effect at the level of the plasma membrane of the target micro-organism (Bohlmann, 1994). The families PR-15, -16 and 17 have been added recently. PR-15 and -16 are typical of monocots and comprise families of germin-like oxalate oxidases and oxalate oxidase-like proteins with superoxide dismutase activity (Berner & Bernia, 2001). These proteins generate hydrogen peroxide that can be toxic to different types of attackers or could directly or indirectly stimulate plant-defense responses (Donaldson *et al.,* 2001). PR-17 proteins have been found as an additional family of

PRs in infected tobacco, wheat and barley and contain sequences resembling the active site of zinc metalloproteinases (Christensen *et al.,* 2002). A putative novel family (PR-18) comprises fungus- and salicylic acid (SA)-inducible carbohydrate oxidases, as exemplified by proteins with hydrogen peroxide-generating and antimicrobial properties from sunflower (Custers *et al.,* 2004).

1.7.2 Osmotin

Osmotins of the same group display similarity to TL proteins (Singh *et al.,* 1987). Osmotin is a basic 24-kDa pathogenesis-related (PR) protein that accumulates in salinity and desiccation-adapted tobacco cells (Singh *et al*., 1985; Singh *et al*., 1987). The expression pattern of this gene indicates that its transcription can be activated by several factors, including NaCl, desiccation, ethylene, wounding, abscisic acid, tobacco mosaic virus, fungi, and UV light (Stintzi *et al.,* 1991). However, in nonhypersensitive-responding genotypes the protein mainly accumulates in response to osmotic stress (LaRosa *et al*., 1992). It has been demonstrated that osmotin has antifungal activity against a variety of fungi, including *Phytophthora infestans*, *Candida albicans*, *Neurospora crassa*, and *Trichoderma reesei* (Woloshuk *et al*., 1991; Vigers *et al*., 1992). Recent reports also indicate that not only tobacco osmotin but also osmotin-like proteins from several plant species have similar inhibitory effects on fungal pathogens and may function as plant defense proteins (Vigers *et al*., 1991).

1.7.3 Protease Inhibitors (PIs)

Plants have numerous PIs [e.g. 41 in Arabidopsis and 52 in rice (http://merops.sanger.ac.uk)] which are widely distributed in plant organs to establish a balance between endogenous proteolysis and protein synthesis (Rawlings *et al.,* 2006). Moreover, they can also function as suppressors of exogenous proteolytic enzymes of phytopathogens, including insects, nematodes, bacteria and fungi (Jackson & Tailor, 1996; Valueva & Mosolov, 2004). PIs are also potentially

important in developing genetically modified crops with elevated resistance to biotic stresses, particularly insects (Reeck *et al.,* 1997; Ferry *et al.,* 2005). PIs can be expressed constitutively or be induced on pest or pathogen attack (Bolter & Jongsma, 1997). In phytophagous insects, PIs either bind to digestive proteases, impairing protein digestion in the gut (Broadway & Duffey, 1986), or may affect moulting and non-digestive enzyme regulation (Faktor & Raviv, 1997). The only concern regarding the use of transgenic plants expressing PIs is the possibility of endangering the predator population involved in the biological control of pests and useful insects, such as honey bees (*Apis mellifera L.*). Ferry *et al.*, (2005) demonstrated that feeding the predator insects with prey that fed on transgenic plants expressing PIs had no detrimental effect on predator biology. However, feeding nurse honey bees with a pollen diet containing at least 1 % soybean trypsin inhibitor can potentially endanger colony growth and maintenance (Sagili *et al.,* 2005).

The co-expression of PIs also appears to be useful during the heterologous expression and isolation of recombinant proteins (Rivard *et al.,* 2006; Kim *et al.,* 2007) and antibodies (Komarnytsky *et al.,* 2006). Constitutive co-secretion of PIs can reduce the extracellular protease activity, which has useful merits in increasing the yields of complex therapeutic proteins secreted from plant tissues (Komarnytsky *et al.,* 2006).

1.7.4 Cellular Localization of PRs

In leaves, PRs appear to be present both in epidermal and mesophyll cells, as well as in the vascular bundles. For example, in response to infection by *Phytophthora infestans*, potato accumulates PR-1b in the vicinity of the successfully colonized leaf area and of the epidermal cell layer in particular. Additional locations within infected leaves were stomatal guard cells, glandular trichomes, crystal idioblasts and the vascular bundles (Hoegen *et al.,* 2002).

Many PR proteins are synthesized with an N-terminal signal peptide determining translocation into the ER, followed by secretion into the apoplast. These proteins accumulate extracellularly and can be collected easily in intercellular washing fluid. PR-type proteins have been collected from xylem fluid of tomato, broccoli, rape, pumpkin, and cucumber (Buhtz *et al.*, 2004; Kehr *et al.*, 2005; Rep *et al.*, 2003; Rep *et al.*, 2002) and from the guttation fluid of barley leaves (Grunwald *et al.*, 2003), suggesting that secretion into the veins entails uptake and transport in the transpiration stream. Other proteins have additional extensions specifying deposition into the vacuole. PR-10-type proteins are the only family of which all members seem to be cytoplasmic (van Loon *et al.*, 2006).

1.7.5 Functional Roles of PRs other than Biotic Stress

The occurrence of homologous PRs as small multigene families in various species belonging to different plant families, their tissue-specific expression during development and consistent localization in the apoplast as well as in the vacuolar compartment, and their differential induction by endogenous and exogenous signalling compounds, suggest that PRs have important functions extending beyond a role in adaptation to biotic stress conditions (van Loon & van Strien, 1999).

Secreted PRs accumulate in senescing leaves of some species and in ripening fruits (Buchanan-Wollanston *et al.*, 2003; Davoine *et al.*, 2001; Quirino *et al.*, 2000), as well as in the medium of cell suspension cultures in the absence of visible necrosis (Broekaert *et al.*, 2000; Takeda *et al.*, 1990), indicating that at least some PRs are also produced under specific physiological conditions.

Abiotic stresses can also elicit pathogenesis-related protein induction, as in the case of osmotic stress, cold stress, or wounding (Broekaert *et al*., 2000). Typically, tobacco osmotin is induced in leaves, stems and roots by drought, high salt, or abscisic acid, as well as in leaves by wounding or UV light. Several reports show that apoplastic PR-1-type proteins, chitinases, glucanases, thaumatin-like proteins, thionins, and lipid-transfer proteins are induced during cold stress. Hon *et al.,* (1995) have shown that PR-2, -3, and -5 proteins have an antifreeze activity on cold hardening in rye. The same proteins accumulated in response to cold, short daylength, and dehydration. The chitinase and glucanase-like proteins, have both enzymatic and antifreeze activities (Atici & Nalbantoglu, 2003; Griffith & Yaish, 2004).

Other abiotic stresses, such as heavy metal toxicity, are likely to induce PRs as a result of the cell-damaging action of the stimulus (van Loon *et al.,* 2006). When testing for the compounds in various pollen and latex that are responsible for allergic reactions in humans, several types of PR-like proteins were found to be responsible. The major pollen allergen from birch, *Bet v 1* (Bufe *et al.,* 1996), belongs to the PR-10 family, and is induced by both biotic and abiotic stress conditions in various plant tissues. PR-10-type proteins are widespread in plants and PR-10-type allergens are present in many foods such as fruits and vegetables. Other pathogenesis-related food allergens belong to the PR-2, -3, -4, -5, -8, -12, -14, and -15 families (Hoffmann-Sommergruber, 2000, Wu *et al.,* 2001).

Several PRs are expressed in cultured cells (Takeda *et al*., 1990) or upon the transition of plants to flowering (Neale *et al*., 1990), suggestive of a developmental role. PR-2-like, PR-3-like and PR-5-like proteins accumulate in the apoplast of winter rye tissues during cold acclimation and exhibit antifreeze activity (Antikainen *et al*., 1996). Basic PR-5 proteins (osmotin) are induced in tobacco and tomato in response to osmotic response.

1.7.6 AtPNP-A as a PR Protein

Although AtPNP-A is yet to be proven to be induced at the protein level in response to pathogens, elevated protein levels have been shown as a result of abiotic stresses (Rafudeen *et al.*, 2003). In addition, transcription of AtPNP-A is low under control conditions but strongly induced in response to biotic and abiotic stresses and the protein has been identified and isolated from the *A. thaliana* apoplast together with PR-1, PR-2 and PR-5 proteins (Boudart *et al.*, 2004). AtPNP-A has other features characteristic of PR proteins including an N-terminal signal peptide (Van Loon *et al.*, 2006) that directs the molecule into the extracellular space. Further, induction of *AtPNP-A* at the transcript level appear to occur independent of *de novo* protein synthesis characteristic of genes encoding secreted proteins (Wang *et al*., 2006). The evolutionary history of AtPNP-A suggests that PNPs, like the related expansins, derived from ancestral family-45 endoglucanases have lost their expansin activity and have sub-functionalized into extracellular, systemically mobile signalling molecules (Ludidi *et al.*, 2002).

CHAPTER TWO: INTRODUCTION TO THE PRESENT STUDY

2. Introduction

Natriuretic peptides (NP) were first identified in vertebrate species ranging from elasmobranchs and mammals where they play a role in the regulation of salt and water balance (Maryani *et al*., 2001). They were recently discovered in plants with the following evidence to support their presence in plants:

- ^{125}I – labeled rANP binds to isolated leaf microsomes (Gehring *et al.,* 1996)
- rANP promotes stomatal opening in *Tradescantia sp.* in a concentration and conformation dependent manner (Suwastika *et al.,* 2000)
- plant natriuretic peptides cause rapid reversible increases in cGMP (Pharmawati *et al.,* 2001)
- two *Arabidopsis thaliana* irPNP-encoding genes termed *AtPNP-A* and *AtPNP-B* have been identified (Ludidi *et al.,* 2002)

The first indications for NPs in plants came from radioimmunoassays on Florida beauty (*Dracena godseffiana*) (Vesely & Giordano, 1991) where antibodies against the N-terminus (ANP, 1–98), the mid-portion (ANP, 31–67) and the C-terminus (ANP, 99–126) recognized peptides in leaves and stems. AtPNP-A has since been extensively studied (reviewed by Gehring, 1999; Gehring & Irving, 2003) in order to elucidate its mode of action and possible function in plants. These studies clearly indicated a functional role for PNP-A in water homeostasis, with the peptide sequence, 33 - 66 in ANP identified as the section corresponding to the activity of AtPNP-A (Morse *et al.,* 2004).

Plant natriuretic peptides are most closely related to expansins, which are key regulators of cell wall extension (Ludidi *et al*., 2002; Li *et al*., 2003; Kende *et al*., 2004). A significant sequence similarity of *At*PNP-A (AAD08935) and *At*PNP-B (CAB79756) to a functionally uncharacterised transcript from citrus, CjBAp12 (AAD03398) has been observed with all three transcripts containing a signal sequence suggesting that the proteins are secreted (Ludidi *et al*., 2002). AtPNP-A and AtPNP-B, as well as CjBAp12, show sequence similarity with the *N*-terminus of expansins but are considerably shorter lacking the wall binding domain (Ludidi *et al*., 2002; Li *et al*., 2003). However, unlike expansins, PNPs are mobile and seem to be involved in regulating solute and water homeostasis (Gehring & Irving 2003; Maryani *et al*., 2003; Rafudeen *et al*., 2003; Morse *et al*., 2004).

2.1 Problem Statement

Plant natriuretic peptides are a novel group of peptides whose function is currently not well elucidated. Two irPNP-encoding genes have been identified namely; *At*PNP-A and *At*PNP-B. There have been extensive studies on the possible function of PNP-A while there is little if any literature at all about the function of PNP-B. In this study we therefore aim to study the evolutionary development of PNP-B as well as perform preliminary studies to establish a possible functional role in either water homeostasis or in plant defense.

2.2 Hypothesis

The function of PNP-B can be deduced from its sequence similarity with PNP-A, expansins, CjBAp12 and other known plant peptides.

2.3 Aim

The amplification, cloning and sequencing of the coding region of PNP-B from various evolutionary related plants and optimization of the heterologous expression of the peptide for functional and structural studies.

2.4 Objectives

1. Bioinformatics analysis of PNP-B
    a. Search for the PNP-B gene using Entrez protein and BLAST and sequence alignments of all the sequences related to AtPNP-B
    b. Chromosomal location and identification of genes surrounding PNP-B using the rice and Arabidopsis genome sequences.
    c. Promoter and regulatory elements prediction of PNP-B in the rice and Arabidopsis genome.
    d. Subcellular location and signal peptide prediction
    e. Secondary and tertiary structure prediction
2. Expression of the novel protein, PNP-B
    a. Subclone the coding region of the PNP-B gene from selected plants into an expression plasmid
    b. Optimization of heterologous expression of the novel protein, PNP-B in *Escherichia coli*
3. The evolutionary relationship of PNP-B in plants
    a. Obtain DNA sequences of PNP-B genes from various plant genomes
    b. Construct evolutionary trees

CHAPTER THREE: BIOINFORMATICS ANALYSIS OF PNP-B

3. Introduction

3.1 Bioinformatics Analysis

The knowledge of protein structure is generally considered a prerequisite to understanding protein function (Bourne & Weissig, 2003; Bergeron, 2003). Approximately 30 000 protein structures are now known. Most were determined by X-ray crystallography or Nuclear Magnetic Resonance (NMR) (Lesk, 2005). Experimental methods are time consuming, expensive and do not give guaranteed success. However, computational methods are the more viable means of predicting protein structures of novel proteins. Computational methods such as *ab initio*, folding recognition and comparative modeling methods are the most widely used methods in the prediction of the protein structure using amino acid sequences (Bourne & Weissig, 2003; Bergeron, 2003).

With the advent of structural genomics, prediction of biological function from structure has become one of the major goals of structural biology and bioinformatics (Shapiro & Harris, 2000). Assignment of biological function provides a valuable first step toward experimental characterization of cellular and physiological roles of gene products (Bourne & Weissig, 2003).

The identification of similar amino acid sequences is used to infer both the structure and the function of a protein. It is believed that structure and function can be transferred between similar sequences because they have been conserved over long periods of time. This has been confirmed for structure (Chothia & Lesk, 1986), but it is a bit more difficult to justify for function. Above

40% sequence identity, homologous proteins tend to have the same function, but below this threshold, conservation of function falls rapidly (Todd *et al.*, 2001). Even at high levels one must be cautious in inferring function, as some sequence relatives with 35% or more sequence identity can have differing catalytic activities (Bourne & Weissig, 2003).

Protein functions are often conferred by a few conserved residues, which sequence-based methods often fail to detect. These residues will, however, be related in 3D, so a comparison of structural similarities between proteins has the potential to identify functional similarities in nonhomologous proteins (Bourne & Weissig, 2003).

Proteins, like genes do not exist as linear sequences of molecules, but they assume complex, compact 3D shapes. Protein shapes or configurations are characterized as secondary, tertiary, or quaternary. The primary structure, which is the linear sequence of the amino acids, is functionally uninteresting (Bergeron, 2003). Secondary and tertiary structures are mostly related to protein function.

We thus have the following paradigm:

- DNA sequence determines protein sequence
- Protein sequence determines protein structure
- Protein structure determines protein function (Lesk, 2005).

3.2 Protein Structure

Proteins are large molecules which, play a variety of roles in life processes: there are structural proteins (for example, viral coat proteins); proteins that catalyse chemical reactions (the enzymes); transport and storage proteins (haemoglobin); regulatory proteins, including hormones and receptor/signal transduction proteins; proteins that control gene transcription; and proteins involved in recognition, including cell adhesion molecules, and antibodies and other proteins of the immune system (Lesk, 2005). In many cases only a small part of the structure – an active site – is directly functional, the rest existing primarily to create and fix the spatial relationship among the active site residues. Proteins evolve by structural changes, produced by mutations in the amino acid sequence and genetic rearrangements, which bring together different combinations of structural subunits (Lesk, 2005).

Protein molecules have four levels of structural organization: primary (sequence), secondary (local folding), tertiary (overall folding) and quaternary (multichain association) (Matthews *et al.,* 2000).

3.2.1 Primary Structure

The primary structure of a protein is the first level of organization and depicts the sequence of amino acids in its polypeptide chain. The primary structure is dictated by the DNA sequence in the gene for each protein (Matthews *et al.,* 2000). A specific sequence of nucleotides in DNA is transcribed into mRNA, which is read by the ribosome in a process called translation. The sequence of a protein is unique to that protein, and defines the structure and function of the protein. The sequence of a protein can be determined by methods such as Edman degradation or tandem mass spectrometry. Often however, it is read directly from the sequence of the gene using

the genetic code. Post-transcriptional modifications such as disulfide formation, phosphorylations and glycosylations are usually also considered a part of the primary structure, and cannot be read from the gene (Cooper & Hausman, 2007). Most proteins exhibit higher levels of structural organization as well. It is the specific three-dimensional structure of each protein that allows it to function in its particular biological role (Matthews *et al.*, 2000).

3.2.2 Secondary Structure

The secondary structure is the regular arrangement of amino acids within localized regions of the polypeptide. Two types of secondary structure, which were first proposed by Linus Pauling and Robert Pauling in 1951, are particularly common: the α-helix and the β-sheet. Both of these secondary structures are held together by hydrogen bonds between the CO and the NH groups of peptide bonds (Cooper & Hausman, 2007). The part of the backbone that is not in a regular secondary structure is said to be a random coil.

3.2.2.1 α-Helix

An α-helix is formed when a region of a polypeptide chain coils around itself, with the CO group of one peptide bond forming a hydrogen bond with the NH group of a peptide bond located four residues downstream in the linear polypeptide (Cooper & Hausman, 2007). This results in a strong hydrogen bond that has the nearly optimum N ·····O distance of 2.8 Å. Amino acid side chains project outward and down ward from the helix, thereby avoiding steric interference with the polypeptide backbone and with each other. The core of the helix is tightly packed; that is, its atoms are in van der Waals contact. Not all amino acids favor the formation of the α-helix due to steric constraints of the R-groups. Amino acids such as alanine, aspartic acid, glutamic acid, isoleucine, leucine and methionine favor the formation of α-helices, whereas, glycine and proline favor disruption of the helix. This is particularly true for proline since it is a pyrrolidine based

amino acid (HN=) whose structure significantly restricts movement about the peptide bond in which it is present, thereby, interfering with extension of the helix. The disruption of the helix is important as it introduces additional folding of the polypeptide backbone to allow the formation of globular proteins (Voet *et al.,* 1999). The helical twist of the α-helix found in all proteins is right-handed (Nelson & Cox, 2005). The α-helix is a common secondary structure encountered in proteins of the globular class.

3.2.2.2 β-Strands

A β-sheet is formed when two parts of a polypeptide chain lie side by side with hydrogen bonds between them. β-sheets can be formed between several polypeptide strands, which can be oriented either parallel or anti-parallel to each other (Cooper & Hausman, 2007). In parallel sheets adjacent peptide chains proceed in the same direction whereas, in antiparallel sheets adjacent chains are aligned in opposite directions (Voet *et al*., 1999).

3.2.3 Tertiary & Quaternary Structure

Proteins show a great variety of three-dimensional conformations. The functions of proteins depend on their adopting the native 3-dimensional structure (Lesk, 2005). Tertiary structure is the folding of the polypeptide chain as a result of interactions between the side chains of amino acids that lie in different regions of the primary structure. In most proteins, combinations of α-helices and β-sheets, connected by loop regions of the polypeptide chain, fold into compact globular structures called domains, which are the basic units of tertiary structure. The formation of tertiary structure is usually driven by the burial of hydrophobic residues in the interior of the protein but other interactions such as hydrogen bonding, ionic interactions and disulfide bonds can also stabilize the tertiary structure. The tertiary structure encompasses all the noncovalent interactions

that are not considered secondary structure, and is what defines the overall fold of the protein, and is usually indispensable for the function of the protein (Cooper & Hausman, 2007).

The highest level of protein structural organization is the quaternary structure. The quaternary structure results from association of independent tertiary structural units through surface interactions, such as formation of the hemoglobin tetramer from myoglobin-like monomers (Cooper & Hausman, 2007). The individual chains are called subunits. The individual subunits are not necessarily covalently connected, but might be connected by a disulfide bond. Not all proteins have quaternary structure, since they might be functional as monomers. The quaternary structure is stabilized by the same range of interactions as the tertiary structure (Nelson & Cox, 2005).

Proteins with multiple polypeptide chains are oligomeric proteins. Oligomeric proteins can be composed of multiple identical polypeptide chains or multiple distinct polypeptide chains. Proteins with identical subunits are termed homo-oligomers. Proteins containing several distinct polypeptide chains are termed hetero-oligomers. Hemoglobin, the oxygen carrying protein of the blood, contains two α and two β subunits arranged with a quaternary structure in the form, α2β2. Hemoglobin is, therefore, a hetero-oligomeric protein (Voet *et al.,* 1999).

3.3 Transcription in Eukaryotes

Transcription is centrally involved in an array of biological processes, which include growth, development and response to external stimuli. Transcription in eukaryotes is a much more complex process than in prokaryotes (Matthews *et al.,* 2000). Transcription refers to the process whereby the double-stranded deoxyribonucleic acid (DNA) blueprint of a gene is converted into a

single stranded messenger ribonucleic acid (mRNA). In eukaryotes, protein-coding genes are transcribed by the RNA polymerase II transcriptional machinery, which comprises RNA polymerase II and other factors that are required for basal and regulated transcription. Gene expression is controlled, *inter alia,* through: (i) dense packaging of chromatin; (ii) transcription initiation through assembly of RNA polymerase II and binding of transcription factors (TFs) at the core promoter; (iii) enhancer elements; (iv) CpG islands; (v) alternative splicing; (vi) polyadenylation; and (vii) translation initiation (Fickett & Hatzigeorgiou, 1997; Johansson *et al.,* 2003; Lareau *et al.,* 2004; Pedersen *et al.,* 1999).

In order for transcription to begin, the RNA Polymerase II holoenzyme and a number of accessory proteins, called general transcription factors (TFs), must bind to the core promoter that spans the transcription start site (TSS) and comprises the TATA box, which has the sequence TATAA and is located 25 to 30 nucleotides upstream of the TSS, initiation region (Inr) and downstream promoter element (DPE), located approximately 30 base pairs downstream of the TSS (Cooper & Hausman, 2007). In addition, regulatory proteins may bind to the proximal promoter, upstream of the core promoter. Transcription resulting from the binding of only the minimal components to the core promoter is called basal transcription and is uncommon *in vivo* (Pedersen *et al.* 1999).

In contrast to basal transcription, the activated transcription that occurs *in vivo* is tightly regulated by the binding of additional TFs to additional *cis*-regulatory elements. Bound TFs mediate their regulatory effects by interacting with the basal transcription complex via protein-protein interactions. This is made possible by the DNA strand bending back on itself (Werner *et al.,* 2003). Regulatory elements can be found several kilobases (kb) away up- or downstream from

the TSS and include both enhancers, that activate transcription, and silencers, that repress transcription (Pedersen *et al.,* 1999).

The type of TF that binds a DNA sequence will also influence whether transcription will be repressed or stimulated, and to what extent. Some TFs binding upstream of a gene may be nonspecific, whilst others may be specific for the gene (Nikolov & Burley, 1997) and the conditions under which transcription must occur. The binding of TFs to DNA is influenced by the degeneracy of the transcription factor binding site (TFBS) and the availability of the TFs. For example, spatial, temporal and environmental conditions may alter the types of TFs that are available to the cell. A specific combination of TFs bound upstream of a gene may thus modulate transcription in a different manner than that of a different combination of TFs bound upstream of the same gene (Pedersen *et al.,* 1999).

3.4 Aims

- **To identify the biophysical properties (functional and structural) of PNP-B and its relation to other molecules as reported in literature.**
- **To predict the promoter region and the regulatory elements of PNP-B.**

3.5 Materials & Methods

3.5.1 Search for the PNP-B gene

The retrieval of all known previously identified plant natriuretic peptide and plant natriuretic peptide-like genes was achieved through use of the Entrez Protein accessed via the NCBI website (http://www.ncbi.nlm.nih.gov/Entrez) (Altschul *et al.,* 1990). The keywords that were used for

retrieving the sequences from Protein Entrez were ‘plant natriuretic peptide’ and the default settings were used. The sequences were saved as FASTA format in a text file.

3.5.2 Sequence Alignments

Sequence alignments were done using CLUSTALW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) (Larkin *et al*., 2007) using the relevant text files that were obtained in section 3.4.1 for inputting the sequences. Default settings were used in all the alignments.

3.5.3 Chromosomal Location and Surrounding Genes

Genes up stream and downstream of AtPNP-B were accessed from NCBI Sequenceviewer (http://www.ncbi.nlm.nih.gov/projects/sviewer/).

3.5.4 Promoter and Regulatory Elements Prediction

Recognition of promoter elements in PNP-B was achieved by scanning the genomic DNA 1 kb upstream and 200 bp downstream of the transcriptional start site (TSS) with Neural Network Promoter Prediction (NNPP; http://www.fruitfly.org/seq_tools/promoter.html) and Promoter Scan (http://www.cbs.dtu.dk/services/promoter) at default settings.

Recognition of regulatory elements in PNP-B was achieved by scanning the genomic DNA 1 kb upstream and 200 bp downstream of the TSS with PLACE (Plant Cis-acting Regulatory DNA Elements; http://www.dna.affrc.go.jp/PLACE/signalscan.html) according to the method of Higo *et al.,* (1999), Genomatix (http://www.genomatix.de/cgi-bin/matinspector_prof/mat_fam.pl?) and Plant Care (Cis-Acting Regulatory Element)

(http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) according to the method of Lescot *et al.*, (2002) were used at default settings.

3.5.5 Subcellular Location and Signal Peptide Prediction

Subcellular location and signal peptide prediction were done for the PNP-B sequence from *Arabidopsis thaliana* using TargetP (http://www.cbs.dtu.dk/services/TargetP/) as described by Emanuelsson *et al*., (2007). Default settings were used in the subcellular location and signal peptide predictions.

3.5.6 Secondary Structure Prediction

Protein secondary structures of PNP-B (Q9M0C2.1) and PNP-A (Q9ZV52.2), from *Arabidopsis thaliana* and CjBAp12 (Q9ZP41.1) (citrus blight protein) were predicted using PSIpred (http://bioinf.cs.ucl.ac.uk/psipred/) as described by Jones (1999). Default settings were used in all the secondary structure predictions.

3.5.7 Tertiary Structure Prediction

The tertiary structure of PNP-B from *Arabidopsis thaliana* was predicted using SWISS-MODEL (http://swissmodel.expasy.org/) (Arnold *et al*., 2006). Default settings were used for the tertiary structure prediction.

3.6 Results

3.6.1 Chromosomal Location and Surrounding Genes

A PNP-B gene from *Arabidopsis thaliana* (Q9M0C2.1) was identified and targeted for amplification and expression as further detailed. The *Arabidopsis thaliana* PNP-B was termed

AtPNP-B for future reference. It was shown that the encoding gene of AtPNP-B sits on chromosome 4 (NC_003075) in the *Arabidopsis thaliana* genome. On the *Arabidopsis thaliana* genome the first 100 base pairs of the AtPNP-B encoding gene correspond to base pair number 14860491 to base pair 14860591. Base pair 101 to base pair 372 on the AtPNP-B encoding gene from *Arabidopsis thaliana*, corresponds to base pair number 14860728 to base pair 14861000 on chromosome 4 of the *Arabidopsis thaliana* genome. This shows the presence of a predicted 138 base pair intron in the encoding gene of AtPNP-B (www.ncbi.nlm.nih.gov/blast/Blast.cgi).

(a)

```
  1 ATGAGTAAAA GTATTGTGTT TTTTTCTACC GTTCTTGTTT TTCTCTTCTC
 51 TTTCTCATAT GCAACTCCCG GGATTGCAAC TTTTTACACA AGTTACACTC
101 GTAAGAATTT TTGTGTGTTT TACTACCAAA ATTATAAAAC AAATACAAAC
151 AACAAACTTA TCTTATACAT TTTCATATCA ACCTACAGCG AAATGTTATA
201 AAATCATTTT CTTTTATTTC ATGTTTGATA GCATCAGCAT GTTACAGAGG
251 TACTCAAGAA GGAGTGATGA TCGCTGCAGC GAGTGATACA TTATGGGACA
301 ATGGTCGAGT TTGTGGCAAA ATGTTCACCG TGAAATGCAG CGGACCTCGT
351 AACGCCGTGC CTCACCCTTG CACGGGGAAA TCCGTGAAGG TCAAGATCGT
401 TGACCATTGT CCCAGTGGCT GTGCTTCTAC GCTCGATCTT TCTCGAGAAG
451 CTTTTGCTCA GATCGCTAAT CCTGTCGCTG GGATTATTAA CATTGATTAT
501 TTTCCGTAA
```

(b)

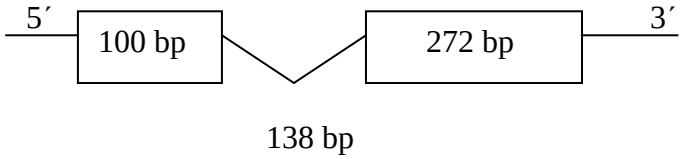


Figure 3.1: Genomic organization of the A*tPNP-B* encoding gene from *Arabidopsis thaliana*. (a) DNA sequence of AtPNP-B with the intron region indicated in red. (b) Schematic representation of the AtPNP-B gene.

Figure 3.1 above shows the genomic organization of AtPNP-B as predicted from the genomic sequence data. The genomic sequence is 510 bp long and contains one predicted 138 bp long intron. The predicted protein is thus expected to consist of 123 amino acids (approximately 13 kDa) and an isoelectric point of 8.4.

The total DNA sequence for chromosome 4 was downloaded from TAIR and the genes surrounding AtPNP-B were identified. It was shown that AtPNP-B is flanked on either side by the zinc finger (C3HC4-type RING finger) family protein with a further heat shock protein and cyclic nucleotide gated channel protein upstream of the 5' zinc finger, as shown in figure 3.2 below.

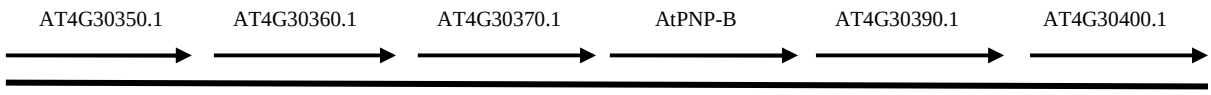


Figure 3.2: Schematic diagram of the genes surrounding AtPNP-B. AT4G30350.1 – heat shock protein-related; AT4G30360.1 – cyclic nucleotide gated channel family; AT4G30370.1 - zinc finger (C3HC4-type RING finger) family protein; AT4G390.1 – Unknown protein; AT4G30400.1 - zinc finger (C3HC4-type RING finger) family protein.

3.6.2 Transcriptional Signals and Regulatory Elements

A sequence ['putative blight-associated protein p12 precursor (oryza sativa japonica group)'; BAD2245.1] from the rice genome was also identified using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) for comparison of promoter and regulatory elements with AtPNP-B. The rice sequence was aligned with both AtPNP-A and AtPNP-B (shown in figure 3.3 above) separately to determine which of the two sequences (AtPNP-A or AtPNP-B), had the highest sequence similarity with the sequence and therefore confirm that it is a PNP-B gene. It was observed that the rice sequence had more sequence similarity with AtPNP-B (54% sequence similarity compared with 38% for AtPNP-A). The *Oryza sativa* PNP-B was termed OsJPNP-B for future reference.

(a)

```
OsJPNP-B        MAKVITSVVIAAVVALAMVSLVAADSGTATFYTPPYTPSACYGFEDQGTMIAAASDVFWN 60
AtPNP-B         MSK---SIVFFSTVLVFLFSFSYATPGIATFYTS-YTP--CYRGTQEGVMIAAASDTLWD 54
                *:*   *:*: :.* : :.*:  * .* *****. ***  **   ::*.*******.:*:

OsJPNP-B        GGAACGQQYVVTCTGPTNQGVPQPCTGQSVTVKIVDHCPSGCAGTIDLSQEAFAIIANPD 120
AtPNP-B         NGRVCGKMFTVKCSGPRN-AVPHPCTGKSVKVKIVDHCPSGCASTLDLSREAFAQIANPV 113
                .* .**: :.*.*:** * .**:****:**.************.*:***:**** ****

OsJPNP-B        AGKVFIDYQQV 131
AtPNP-B         AGIINIDYFP- 123
                ** : ***
```

(b)

```
OsJPNP-B        -MAKVITSVVIAAVVALAMVSLVAADSGTATFYTPPYTPSACYGFEDQGTMIAAASDVFW 59
AtPNP-A         MIKMAVKFVVVMIVFAQILAPIAEAAQGKAVYYDPPYTRSACYGTQ-RETLVVGVKNNLW 59
                 :  .:. **:  *.*  :..:. * .*.*.:* **** ***** : : *::....: :*

OsJPNP-B        NGGAACGQQYVVTCTGPTNQGVPQPCTGQSVTVKIVDHCPSGCAGTIDLSQEAFAIIANP 119
AtPNP-A         QNGRACGRRYRVRCIGATYN-FDRACTGRTVDVKVVDFCREPCNGDLNLSRDAFRVIANT 118
                :.* ***::* * * *.* : . :.***::* **:**.* . * * ::**::** :***.

OsJPNP-B        DAGKVFIDYQQV 131
AtPNP-A         DAGNIRVVYTPI 130
                ***:: : *  :
```

Figure 3.3 Sequence Alignments (a) Alignment of AtPNP-B and OsJPNP-B (b) Alignment of AtPNP-A and OsJPNP-B.

The transcription start site was found -31 bp upstream of the AtPNP-B start codon (ATG) and -109 upstream of the OsJPNP-B start codon (ATG). The core promoter region was predicted to be in the region -71 to -21 bp upstream of the start codon for AtPNP-B and in the region -149 to -99 upstream of the start codon for OsJPNP-B. No downstream promoter element (DPE) was found in both AtPNP-B and OsJPNP-B but three initiator (INRNTPSADB) sites were found for AtPNP-B and one initiator (INRNTPSADB) site for OsJPNP-B. A putative polyadenylation signal is found 119 bp downstream of the stop codon for AtPNP-B and 83 bp downstream of the stop codon for OsJPNP-B.

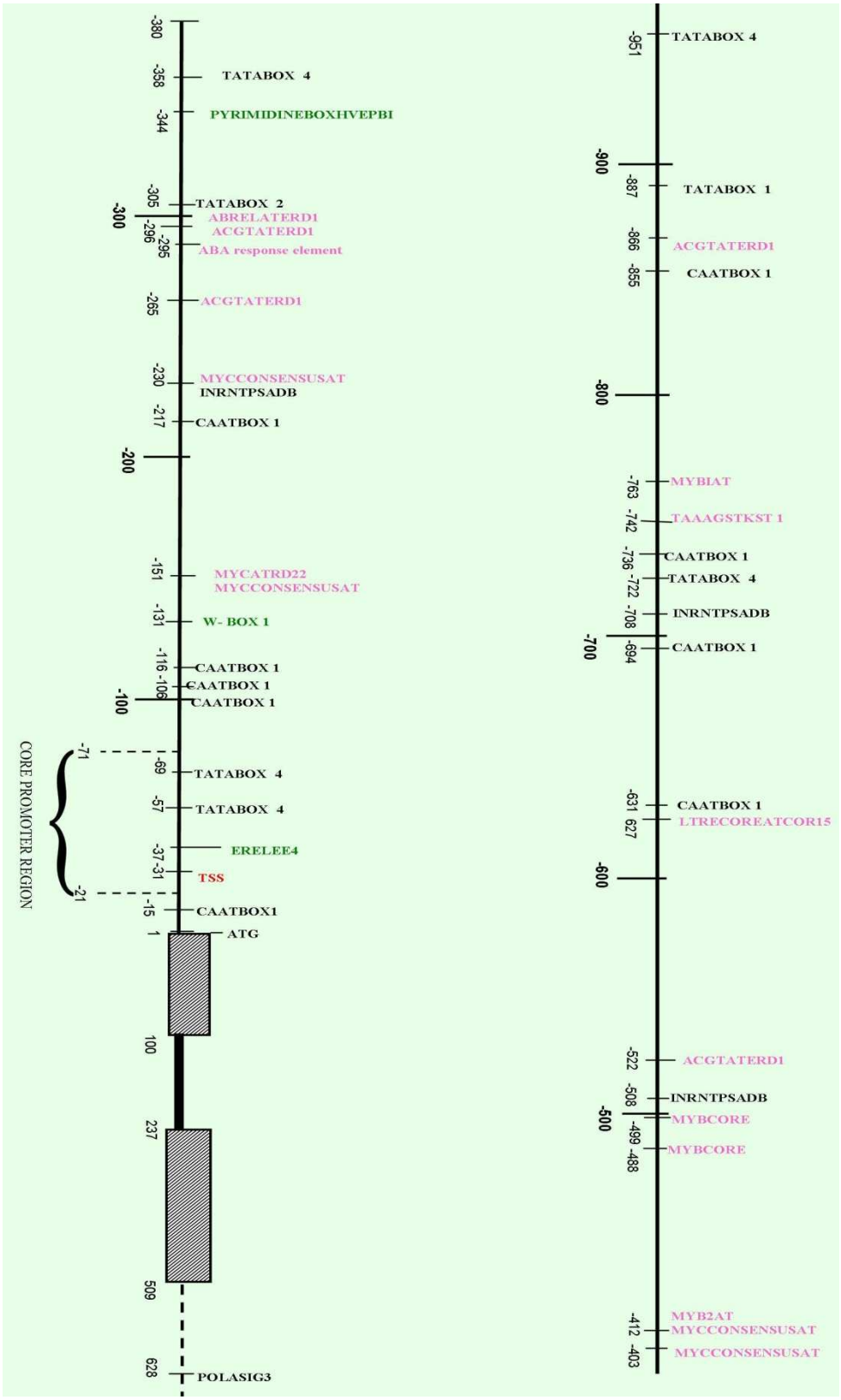


Figure 3.4: Schematic diagram showing the *cis* regulatory elements for the AtPNP-B gene. The abiotic TFBS are shown in pink while the biotic ones are shown in green.

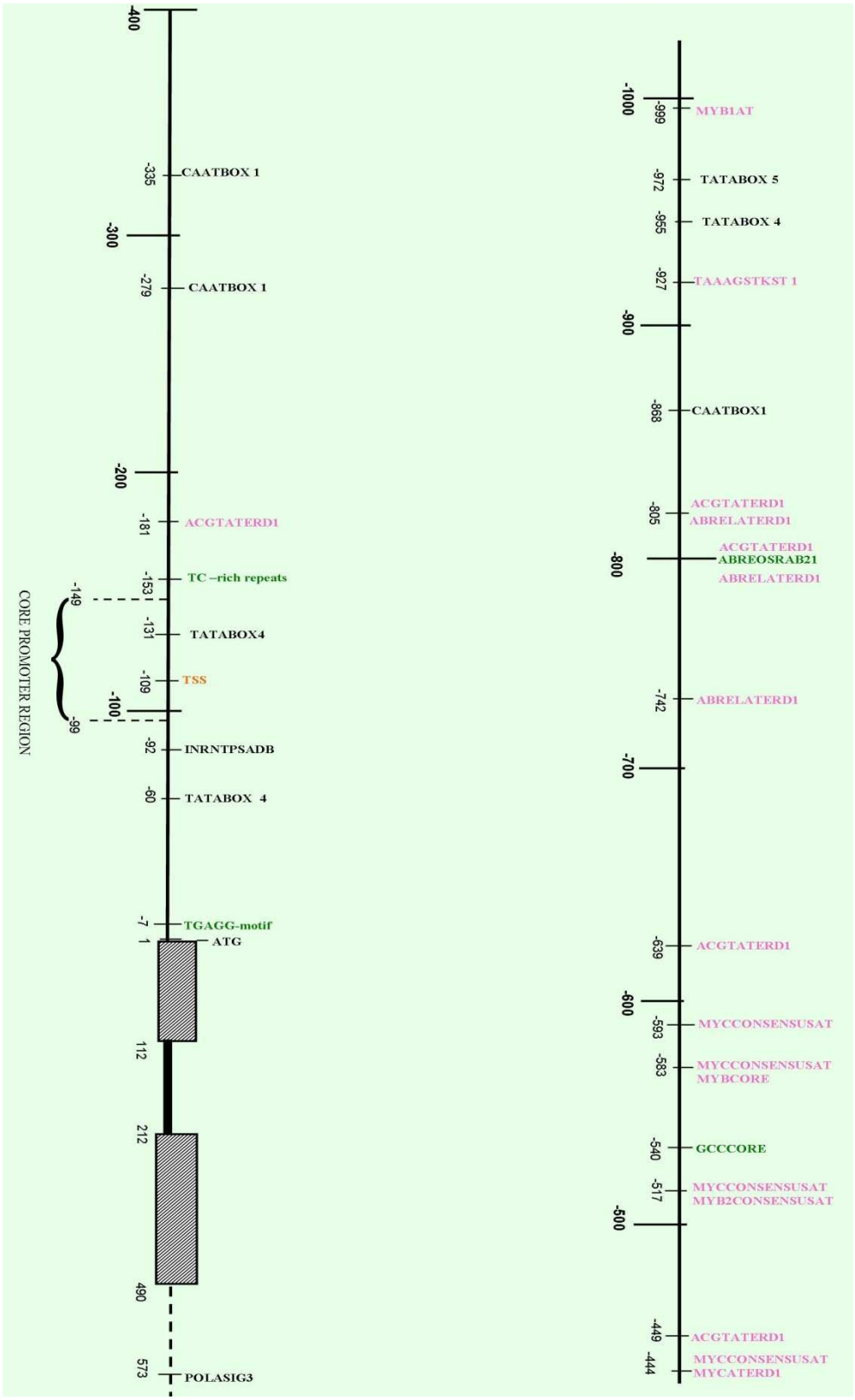


Figure 3.5: Schematic diagram showing the *cis* regulatory elements for the OsJPNP-B gene. The abiotic TFBS are shown in pink while the biotic ones are shown in green.

A summary of the predicted transcription factor binding sites (TFBS) are found in Appendix 2, and represented schematically in figure 3.4 for AtPNP-B and in figure 3.5 for OsJPNP-B. For AtPNP-B, most of the TFBS were shown to be related to water homeostasis in particular dehydration responsiveness and water stress; ABRELATERD1 (Nakashima *et al.,* 2006), ACGTATERD1 (Simpson *et al.*, 2003), MYB1AT (Abe *et al.*, 2003), MYB2AT (Urao *et al.*, 1993), MYCCONSENSUSAT (Abe *et al.*, 2003), MYBCORE (Solano *et al.*, 1995), TAAAGSTKST1 (Plesch *et al.*, 2001) and MYCATRD22 (Abe *et al.*, 1997). There were also other TFBS that were related to abiotic stresses such as low temperature response elements (LTRECOREATCOR15) (Baker *et al.*, 1994).

The AtPNP-B gene was also shown to have predicted TFBS that are related to biotic stresses. These TFBS included, EREELEE4 (ethylene responsive element; Rawat *et al.,* 2005) and the W-Box 1 (Yu *et al.*, 2001), which is a TFBS that has a function in wounding and pathogen response.

The promoters that were predicted to regulate the OsJPNP-B gene were mostly related to water homeostasis in particular dehydration and water stress; ACGTATERD1 (Simpson *et al.*, 2003), MYCATERD1 (Simpson *et al.*, 2003; Tran *et al.*, 2004), MYCCONSENSUSAT (Abe *et al.*, 2003), MYB2CONSENSUSAT (Abe *et al.*, 2003), MYBCORE (Solano *et al.*, 1995), ABRELATERD1 (Nakashima *et al.*, 2006), TAAAGSTKST1 (Plesch *et al.*, 2001) and MYB1AT (Abe *et al.*, 2003). The OsJPNP-B gene was also shown to have predicted TFBS that are related to biotic stresses. These TFBS included, TGAGG-motif (methyl jasmonate responsiveness), TC-rich repeats (defense and stress responsiveness), GCCCORE (ethylene and jasmonate responsiveness; Brown *et al.*, 2003) and ABREOSRAB21 (abscisic responsive element; Busk & Pages, 1998).

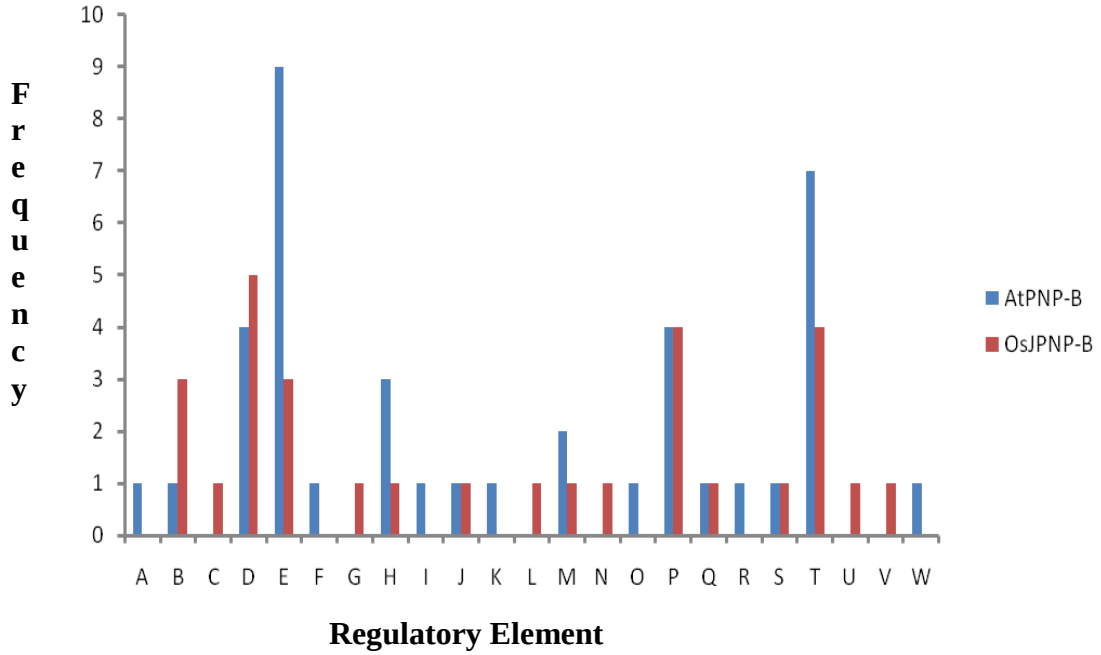


Figure 3.6: Comparative analysis of cis-elements identified in silico for AtPNP-B and OsJPNP-B. The query sequence for each gene was 1 kb upstream of the gene and 200 bp downstream. Elements identified: A – ABA-response element; B – ABRELATERD1; C – ABREOSRAB21; D – ACGTATERD1; E – CAATBOX1; F – ERELEE4; G – GCCCORE; H – INRNTPSADB; I – LTRECOREATCOR15; J – MYB1AT; K – MYB2AT; L – MYB2CONSENSUSAT; M – MYBCORE; N – MYCATERD1; O – MYCATRD22; P – MYCCONSENSUSAT; Q – POLASIG3; R – PYRIMIDINEBOXHVEPB1; S – TAAAGSTKST1; T – TATABOX (1, 2, 4, 5); U – TC-rich repeats; V – TGAGG-motif; W – W-BOX1

The frequency and occurrence of the TFBS for AtPNP-B and OsJPNP-B were compared as shown in figure 3.6 above. The comparison shows that both AtPNP-B and OsJPNP-B have TFBS that relate to both water and biotic stress. In both genes the TFBS that relate to water stress and homeostasis out-number the ones that relate to biotic stress approximately four fold. The TFBS that were present for both AtPNP-B and OsJPNP-B were ABRELATERD1, ACGTATERD1, MYB1AT, MYBCORE, MYCCONSENSUSAT and TAAAGSTKST1. The TFBS that were common for both AtPNP-B and OsJPNP-B were the ones that are related to water stress and dehydration responsiveness. The core promoter elements (CAATBOX and TATABOX) were

also identified for both AtPNP-B and OsJPNP-B. However there were a greater number of core promoter elements for AtPNP-B than OsJPNP-B.

The TFBS that were identified for AtPNP-B only were ABA-response element, ERELEE4, LTRECOREATCOR15, MYB2AT, MYCATRD22, PYRIMIDINEBOXHVEPB1 and W-BOX1. The cis-regulatory elements that were identified for OsJPNP-B were ABREOSRAB21, GCCCORE, MYB2CONSENSUSAT, MYCATERD1, TC-rich repeats and TGACG-motif.

3.6.3 Sequence Alignments

Figure 3.7 below shows the alignments of AtPNP-B with the different molecules that have been shown to be related to it. Figure 3.7a below shows the alignment of AtPNP-B with a second plant natriuretic peptide sequence, from *Arabidopsis thaliana*, denoted as AtPNP-A (Q9ZV52.2) that shows 37% sequence similarity with AtPNP-B. As shown in Figure. 3.7b, AtPNP-B shows a strong sequence similarity to CjBAp12 (Q9ZP41.1), a blight-induced protein of unknown function, isolated from citrus, *Citrus jambhiri* (Ceccardi *et al.*, 1998). AtPNP-B is annotated as a blight-associated protein homolog (BAPH, accession no. CAB79756), since it shows a high degree (54% identity) of sequence similarity with CjBAp12 (accession no. AAD03398). The cysteine residues in both AtPNP-A and AtPNP-B (Figure. 3.7a below) are aligned suggesting a conserved protein folding pattern. While CjBAp12 contains an additional cysteine in its putative signal peptide, all other cysteine residues align with the AtPNP-A and AtPNP-B (Ludidi *et al.*, 2002) (Figure. 3.7c below).

(a)

```
AtPNP-B         -MSKSIVFFSTVLVFLFSFS---YATPGIATFYT---SYTPCYRGTQEGVMIAAASDTLW 53
AtPNP-A         MIKMAVKFVVVMIVFAQILAPIAEAAQGKAVYYDPPYTRSACY-GTQRETLVVGVKNNLW 59
                 :. :: *. .::**   ::    *: * *.:*    : :.** ***. .::....:.**

AtPNP-B         DNGRVCGKMFTVKCSGPRNAVPHPCTGKSVKVKIVDHCPSGCASTLDLSREAFAQIANPV 113
AtPNP-A         QNGRACGRRYRVRCIGATYNFDRACTGRTVDVKVVDFCREPCNGDLNLSRDAFRVIANTD 119
                :***.**: : *:* *.   . :.***::*.**:**.* . * . *:***:**  ***.

AtPNP-B         AGIINIDYFP- 123
AtPNP-A         AGNIRVVYTPI 130
                ** *.: * *
```

(b)

```
AtPNP-B         ---MSKSIVFFSTVLVFLFSFSYATPGIATFYTS-YTP--CYRGTQEGVMIAAASDTLWD 54
CjBAp12         MGVGTKVLVITTMAICLISSAAYASEGTATFYTPPYVPSACNGYKNDGVMIAAASYAIWN 60
                    :* :*: : .: :: * :**: * *****. *.*  *   .::******** ::*:

AtPNP-B         NGRVCGKMFTVKCSGPRN-AVPHPCTGKSVKVKIVDHCPSGCASTLDLSREAFAQIANPV 113
CjBAp12         NGAVCNKSFRVKCTGATNQGTPHPCRGGSVLVKIVDLCPAGCQATIDLSQEAFSQIANPD 120
                ** **.* * ***:*. * ..**** * ** ***** **:** :*:***:***:*****

AtPNP-B         AGIINIDYFP- 123
CjBAp12         AGKIKIEFNQA 131
                ** *:*::
```

(c)

```
AtPNP-B         ----MSKSIVFFSTVLVFLFSFSYATPGIATFYTS-YTP--CYRGTQEGVMIAAASDTLW 53
CjBAp12         -MGVGTKVLVITTMAICLISSAAYASEGTATFYTPPYVPSACNGYKNDGVMIAAASYAIW 59
AtPNP-A         MIKMAVKFVVVMIVFAQILAPIAEAAQGKAVYYDPPYTRSACYGTQRETLVVGVKNN-LW 59
                      * :*.      :: . : *: * *.:* . *.   *    .: :::.. .  :*

AtPNP-B         DNGRVCGKMFTVKCSGPRN-AVPHPCTGKSVKVKIVDHCPSGCASTLDLSREAFAQIANP 112
CjBAp12         NNGAVCNKSFRVKCTGATNQGTPHPCRGGSVLVKIVDLCPAGCQATIDLSQEAFSQIANP 119
AtPNP-A         QNGRACGRRYRVRCIGATY-NFDRACTGRTVDVKVVDFCREPCNGDLNLSRDAFRVIANT 118
                :** .*.: : *:* *.      :.* * :* **:** *   * . ::**::**  ***.

AtPNP-B         VAGIINIDYFP- 123
CjBAp12         DAGKIKIEFNQA 131
AtPNP-A         DAGNIRVVYTPI 130
                 ** *.: :
```

(d)

```
AtPNP-B         P-----CYRGTQEGVMIAAASDTLWDNGRVCGK-MFTVK 33
AtPNP-A         PYTRSACY-GTQRETLVVGVKNNLWQNGRACGR-RY--- 34
ANP             SLRRSSCFGGRMDR---IGAQSGL---G--CNSFRY--- 28
                 .     *: *        .... *   *  *.   :
```

Figure 3.7: Protein sequence alignments. **(a)** Alignment of AtPNP-B and AtPNP-A from *Arabidopsis thaliana* **(b)** Alignment of AtPNP-B and CjBAp12 **(c)** Alignment of AtPNP-B, AtPNP-A and CjBAp12 **(d)** Alignment of AtPNP-B, AtPNP-A (33 - 66) and ANP (99 -126); asterisks (*) identify identical amino acids, colons (:) are conservative amino acid substitutions and dots (.) are semi-conservative amino acid substitutions.

IrPNPs were first isolated through immunoaffinity purification with an antibody directed against the biologically active C-terminus of human α-ANP (α-hANP; NP_006163.1). Previous investigations have suggested that the portion 33 – 66 of AtPNP-A has similarity to human ANP

(Morse *et al.*, 2004). Thus, there is a possibility that the portion 33 - 66 in PNP-B may show some similarity to ANP. The alignment of AtPNP-A, AtPNP-B (33 - 66) and the C-terminus of α-hANP (99 - 126) (Figure. 3.7d above) shows considerable similarity between the two molecules. There is also the conservation of the two cysteine groups in all the three molecules. This might suggest a possibility of disulphide bonds which are critical in maintaining secondary structures.

3.6.4 Secondary Structure, Subcellular Location and Signal Peptide Prediction

A secondary structure prediction program using PSIpred (Jones, 1999) was applied to test for the presence of coil, β-sheet and helix domains. The comparison of AtPNP-B with both AtPNP-A and CjBAp12 shows near congruence in predicted secondary structure as shown in Figure. 3.8 below and might imply common, albeit unknown functionality.

Table 3.1: TargetP v1.1 prediction results (Emmanuelsson *et al*., 2007)

Name	**Len**	**cTP**	**mTP**	**SP**	**other**	**Loc**	**RC**
AtPNP-B	123	0.004	0.028	0.991	0.031	S	1

C Chloroplast, sequence contains **cTP**, a chloroplast transit peptide;

M Mitochondrion, sequence contains **mTP**, a mitochondrial targeting peptide;

S Secretory pathway, sequence contains **SP**, a signal peptide;

RC – "Reliability Coefficient" – a measure of how confident TargetP is in each prediction.

(a)

(b)

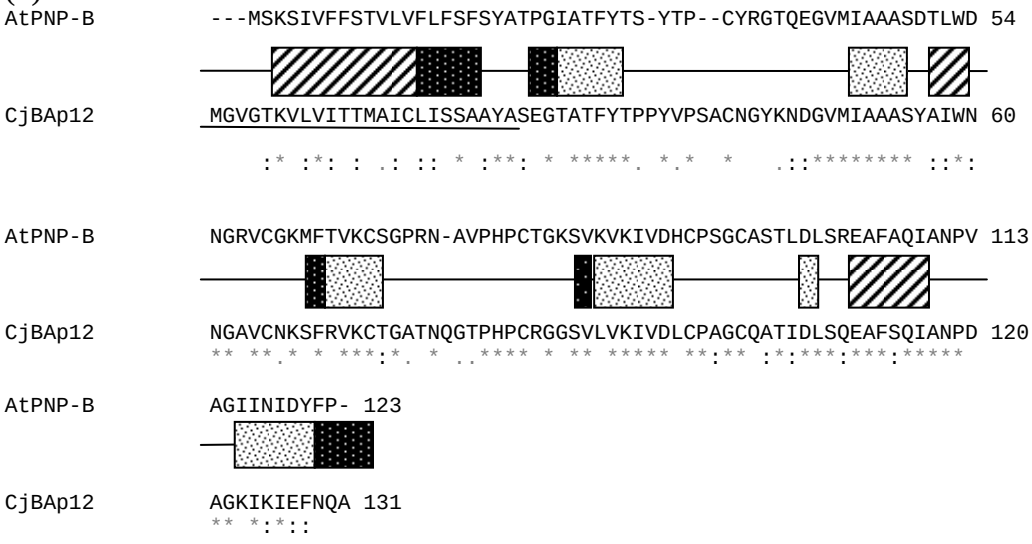


Figure 3.8: Comparison of predicted secondary structures of AtPNP-B with (a) AtPNP-A and (b) CjBAp12 from *Citrus jambhiri*; asterisks (*) identify identical amino acids, colons (:) are conservative amino acid substitutions and dots (.) are semi-conservative amino acid substitutions and the underlined sequence delineates the hydrophobic signal peptide.

Table 3.1 above shows that PNP-B is a secretory protein due to the higher probability of 0.991 of the protein following a secretory pathway. Furthermore, the Reliability Coefficient (RC) obtained for AtPNP-B was 1. Reliability Coefficient is a measure of the reliability of the prediction. The RC ranges from 1 (very reliable prediction; virtually no false positives detected in the TargetP test set) to 5 (not reliable prediction; many false positives detected). The RC is based on the difference between the highest and the second highest TargetP output scores: if this difference is larger than 0.8, then RC is 1; if it is between 0.6 and 0.8, then RC is 2, and so on (Emmanuelsson *et al*., 2007). The prediction of PNP-B to be a secretory protein is consistent with the prediction of the protein to have a signal peptide which is cleaved between amino acid 21 and 22 as shown in Figure 3.9 below.

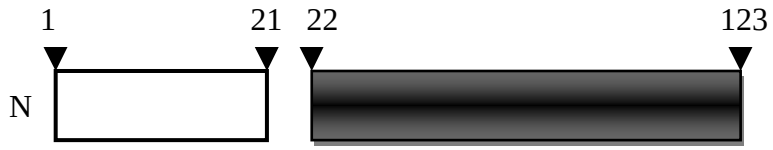


Figure 3.9: Domain organization of AtPNP-B. The open square is the signal peptide and the filled region represents the mature PNP-B peptide.

3.6.5 Tertiary Structure Prediction

The fold model of AtPNP-B was generated using the AtPNP-B sequence (1 - 123) and it was based on the crystal structure of the N-terminal of Ph1 P 1 Timothy Grass Pollen Allergen (accession number P43213). The template has also been used by Ludidi *et al.,* (2004) and Wang *et al.,* (2007) in the prediction of the plant natriuretic peptide tertiary structures. Template selection, alignment and model building were done completely automated by the server (Schwede *et al*., 2003). The template was selected from the SWISS-MODEL template library (SMTL), which is extracted from the Protein Data Bank (Westbrook *et al.,* 2003) and the template had a sequence identity of more than 25% with the submitted AtPNP-B target sequence. The basic

common fold for these molecules is a double psi β barrel structure (Figure 3.10 below) where a six-stranded β barrel assumes a pseudo-twofold axes in which the parallel strands form two psi structures (Castillo *et al.*, 1999).

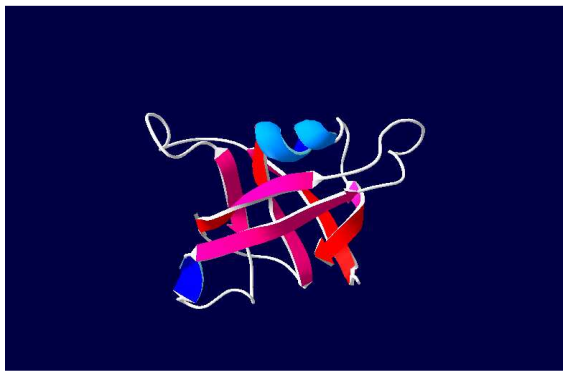


Figure 3.10: Modelled fold of AtPNP-B showing the six stranded double-psi β barrel (DPBB) structure. The α-helices are indicated by the blue twisted ribbons and the β strands as pink ribbons and the two protruding psi loops are marked. The model was generated using SWISS-MODEL (Arnold *et al.*, 2006) using the N-terminal of Ph1 P 1 Timothy Grass Pollen Allergen (accession number P43213) as a template.

3.7 Discussion

Plant natriuretic peptides represent a novel class of small proteins that are closely related to expansins, which are key regulators of cell wall extension (Ludidi *et al.*, 2002; Li *et al.*, 2003; Kende *et al.*, 2004). Thus, homology searches using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) identified mostly expansins as the protein with significant sequence similarity to the AtPNP-B sequence. However, significant sequence similarity of AtPNP-B with AtPNP-A, ANP and CjBAp12 was also observed.

Sequence alignment is fundamental to inferring homology (common ancestry) and function. If the sequence of a protein or other molecule significantly matches the sequence of a protein with a known structure and function, then the molecules may share structure and function (Bergeron,

2003). Multiple sequence alignment, in which three or more sequences must be aligned, is useful in finding conserved regulatory patterns in nucleotide sequences and for identifying structural and functional domains in protein families (Bergeron, 2003).

Sequence alignments of AtPNP-B with ANP, AtPNP-A and CjBAp12 were done and significant sequence similarity was observed. A blight associated protein, CjBAp12 shows the highest sequence similarity (54%) with AtPNP-B thus giving PNP-B a possible function in pathogenesis response. Citrus blight causes zinc deficiency in leaves and accumulation of zinc in bark and wood up to four times the levels found in healthy trees (Albrigo & Young, 1981). Xylem vessels become plugged, reducing water flow throughout the tree, resulting in wilting, reduction in leaf and fruit size and eventual limb dieback (Cohen, 1974). Affected trees show symptoms of water stress due to the xylem dysfunction (Syvertsen *et al*., 1980).

CjBAp12 has a high sequence similarity to AtPNP-B but CjBAp12 is not found in healthy plants while AtPNP-B has been amplified from RNA isolated from healthy, unstressed plants. The presence of PNP-B in healthy plants might suggest that the peptide is not pathogen responsive but this does not preclude a stress-induced up-regulation (Maryani *et al.*, 2001). This will help to support the fact that due to the high similarity of AtPNP-B to CjBAp12, AtPNP-B could have a pathogenesis-related response in infections where it counteracts the disturbances of cation and water transport or directly towards the pathogens that cause the infection. It can therefore, be speculated that PNP-B and CjBAp12 either work together or independently towards the goal of counteracting disturbances of cation and water transport in plants.

To further support the possibility of the function of PNP-B in pathogenesis-related response, the results from the TFBS prediction showed the presence of the W-BOX and the ERELEE4

(ethylene responsive element) in the region upstream of the AtPNP-B coding region. The W-BOX is recognized specifically by salicylic acid – induced WRKY DNA binding proteins (Yu *et al.,* 2001). This gives PNP-B a possible function in plant systemic acquired resistance (SAR) defense responses which may involve the modification of cellular ion and water homeostasis during stress responses. SAR can be triggered by exposing the plant to virulent, avirulent, and non-pathogenic microbes, or artificially with chemicals such as salicylic acid, 2,6-dichloro-isonicotinic acid (INA) or benzo (1,2,3) thiadiazole-7-carbothionic S-methyl ester (BTH) (Sticher *et al.,* 1997). SAR results in the coordinated accumulation of pathogenesis-related proteins (and transcripts) and salicylic acid throughout the plant (Cameron *et al.,* 1994).

The prediction of a possible function of AtPNP-B in the SAR pathway is consistent with the work of Meier *et al.,* (2008). Meier *et al.,* (2008) have shown that the expression of AtPNP-A, the other PNP sequence from *Arabidopsis thaliana* that shows similarity (34%) with AtPNP-B, is significantly correlated with that of genes involved in the SAR defense response pathway in response to various biotic and abiotic stimuli.

The presence of ERELEE4 (ethylene responsive element; Rawat *et al.,* 2005) as one of the TFBS of AtPNP-B also shows that PNP-B could be involved in induced systemic resistance (ISR). Unlike SAR, ISR does not involve the accumulation of pathogenesis-related proteins or salicylic acid (Pieterse *et al.,* 1996), but instead, relies on pathways regulated by jasmonate and ethylene (Yan *et al.,* 2002). OsJPNP-B also contains two cis-acting elements involved in jasmonate responsiveness namely; the TGACG-motif and GCCCORE, further supporting the possible function of PNP-B in ISR. The GCC-box is found in many pathogen-responsive genes such as PDF1.2, Thi2.1 and PR4 (Brown *et al.,* 2003). GCC-boxes have been shown to function as ethylene-responsive elements (Koyama *et al.,* 2003) and they also appear to play important roles

in regulating jasmonate-responsive gene expression (Brown *et al.,* 2003). The jasmonate-responsive elements are found in OsJPNP-B alone but both AtPNP-B and OsJPNP-B have ethylene-responsive elements. Jasmonate and ethylene are reported to aid in pathogen resistance (Rance *et al.,* 1998). Thus, PNP-B could be involved in providing plants with resistance against fungi and bacteria. TC-rich repeats, cis-acting elements that are involved in defense and stress responsiveness were also predicted upstream of the OsJPNP-B gene.

For both AtPNP-B and OsJPNP-B abscisic acid (ABA) responsive (ABA response element and ABREOSRAB21) cis-acting regulatory elements were predicted. ABA is well known to regulate responses to abiotic factors such as salinity and drought (Zhu, 2002). However, more recently, ABA has also been implicated in the establishment of compatible interactions between fungal pathogens and host plants (McDonald & Cahill, 1999; Mohr & Cahill, 2003). Thus, the presence of this cis-regulatory element in both PNP-B genes implicates a role for PNP-B in plant defense responses which may involve the modification of cellular ion and water homeostasis during stress response.

When AtPNP-B was aligned with AtPNP-A it showed a 34% similarity. This implied that AtPNP-B has a possible role in water homeostasis as has been studied before (Morse *et al*., 2004). Studies have shown that AtPNP-A is an osmotic stress-responsive peptide hence; AtPNP-B might also have a role associated with water homeostasis due to the similarity between the two.

In order to help confirm the possible function of PNP-B in water homeostasis, the transcription factor binding sites (TFBS) were analysed using the *Arabidopsis* (AtPNP-B) and Rice (OsJPNP-B) genomes. The result from the TFBS, section 3.5.2 as shown in figures 3.4 and 3.5 shows that PNP-B is most likely to have a function in water homeostasis as noted by the fact that most of the

TFBS were related to dehydration responsiveness and water stress. The frequencies of the water related TFBS in both AtPNP-B and OsJPNP-B was shown to be equal (17 for both) despite the higher water requirements for rice. This shows us that PNP-B regulation is independent of the extent of water requirements of the plant.

Examples of TFBS that were shown to be related to dehydration responsiveness were ABRELATERD1 (Nakashima *et al.*, 2006), ACGTATERD1 (Simpson *et al.,* 2003), MYB1AT (Abe *et al.,* 2003), MYCARTRD22 (Abe *et al.,* 1997) and MYCCONSENSUSAT (Abe *et al.,* 2003). Other interesting TFBS that were shown to be related to water stress were MYB2AT (Urao *et al.,* 1993) and MYBCORE (Solano *et al.,* 1995) which are the binding sites for ATMYB2. ATMYB2 is involved in the regulation of genes that are responsive to water stress in *Arabidopsis*. Another TFBS, which contains a MYB recognition site that was found downstream of the OsJPNP-B gene, was MYB2CONSENSUSAT. The MYB recognition site is found in the promoters of the dehydration-responsive gene rd22 (Abe *et al.,* 2003).

Another TFBS that was found upstream of both AtPNP-B and OsJPNP-B was TAAAGSTKST1. The TAAAG elements are target sites for trans-acting Dof proteins controlling guard cell-specific gene expression (Plesch *et al.,* 2001). The presence of the TAAAGSTKST1 cis-regulatory element upstream of both AtPNP-B and OsJPNP-B is consistent with the experiments of Gehring *et al.,* (1996) which demonstrated that synthetic rANP can induce stomatal opening in *Tradescantia* sp. in a concentration dependent manner. Biologically active PNP immune-analogues (irPNP) and recombinant AtPNP-A have also been shown to promote stomatal opening in a concentration dependent manner (Pharmawati *et al.,* 1998a, 2001; Maryani *et al.,* 2001; Morse *et al.,* 2004). Therefore, this implies that PNP-B could also promote stomatal opening.

The experiments of Gehring *et al.,* (1996) have shown that Na^+ is required in the medium for activity in animal systems, but it is not the case in plants. Gehring & Irving (2003) suggested that in plants, natriuretic peptides operate on processes other than Na^+ transport, such as K^+ transport or the synthesis of compatible solutes. This is supported by the identification of the TAAAGSTKST1 regulatory element from both AtPNP-B and OsJPNP-B because the TAAAG motif is found in the promoter of the KST1 gene which encodes a K^+ influx channel of guard cells in potato (Plesch *et al.,* 2001). Thus, it can be concluded that PNP-B operates on K^+ transport and has a possible function in promoting stomatal opening.

PNPs are recognized by antibodies against vertebrate atrial natriuretic peptides (ANPs) (Vesley *et al.,* 1993; Billington *et al.,* 1997; Gehring, 1999). Vertebrate ANPs are a highly conserved and well studied group of peptide hormones with an important role in vertebrate salt and water homeostasis (Kourie & Rive, 1999; Kone, 2001; Rafudeen *et al.,* 2003). PNPs and vertebrate ANPs can elicit a number of common biological responses in plants (Gehring, 1999; Gehring & Irving, 2003) and one such response is the promotion of stomatal pore opening that causes increased transpiration and gas exchange. While PNPs and ANPs can be viewed as functional analogues, they are not likely to be related molecules in the evolutionary sense (Ludidi *et al.,* 2002). The weak similarity between AtPNP-B and human ANP is the result of convergent evolution (Ludidi *et al.,* 2002).

The region of AtPNP-A (33 - 66) that has homology to ANP has also been shown to trigger protoplast swelling and stomatal guard cell opening (Morse *et al.,* 2004). When the region 33 – 66 of AtPNP-A and AtPNP-B were aligned with ANP some similarities were found. The sequence alignment showed conservation of the two cysteine residues in the three molecules. This suggests the formation of a disulphide bond which is important in maintaining the function

of the molecule (Morse et al., 2004). This is supported by the experiments of Pharmawati *et al.*, (1998a) who showed that, ANP when linearized loses its biological activity. The experiments of Wang *et al.,* (2007) have also shown that the linearization of AtPNP-A removes its ability to promote increases in protoplast volume. The conservation of the two cysteine residues in the three molecules (AtPNP-A, AtPNP-B and CjBAp12) and the experimental evidence, strongly suggests that the effect of PNPs is not just encoded in the primary amino acid sequence but is dependent on secondary and tertiary structure.

In 1999, the Nobel prize in Physiology or Medicine was awarded to Günther Blobel for the discovery that proteins have intrinsic signals that govern their transport and localization in the cell. The determination of the subcellular localization of a protein is an important first step towards understanding its function (Emanuelsson *et al.,* 2007) therefore, subcellular localization and prediction was one of the bioinformatics tools that was used to determine the localization of PNP-B. AtPNP-B was shown to follow a secretory pathway.

The best known protein ‘‘zip code’’ is the secretory signal peptide (SP), which is found in all the three domains of life. It targets a protein for translocation across the plasma membrane in prokaryotes and across the endoplasmic reticulum (ER) membrane in eukaryotes (von Heijne, 1990). AtPNP-B was found to be a secretory protein and it also has a signal peptide between amino acid 21 and 22. The signal peptide is cleaved off during its translocation across the membrane making PNP-B a translocatable protein. This shows that PNP-B is associated with the membranes and also supports the evidence that several irPNP-dependent processes have been observed in experimental systems that did not contain cell walls such as protoplast or microsomal and plasma membrane vesicles. The processes include: *in vitro* irPNP binding to isolated Tradescantia multiflora leaf microsomes (Suwastika *et al.,* 2000), irPNP-dependent modulation

of plasma-membrane H^+ gradients in potato leaf tissue vesicles (Maryani *et al.*, 2000), increases of cGMP levels in response to irPNP in potato guard cell protoplasts (Pharmawati *et al.,* 2001) and irPNP-dependent volume changes in protoplasts (Maryani *et al.,* 2001). This implies that PNP-B can act directly on the plasma membrane.

AtPNP-A has an important and systemic role in plant growth and homeostasis (Morse *et al.*, 2004). The accumulation of CjBAp12 in the leaf tissues when it is not synthesised there, indicates systemic mobility (Ceccardi *et al.,* 1998). AtPNP-A, CjBAp12 and AtPNP-B all have signal peptides. This suggests that AtPNP-B is a systemically mobile protein due to the similarity that it has to AtPNP-A and CjBAp12.

The tertiary structure of AtPNP-B was modelled based on the crystal structure of the N-terminal of Ph1 P 1 Timothy Grass Pollen Allergen (accession number P43213). Ph1 P 1 Timothy Grass Pollen Allergen is a group-1 pollen allergen. Group-1 pollen allergens of grasses have limited but significant sequence homology to expansins. Thus, the modelling of AtPNP-B using the pollen allergen verifies the similarity of plant natriuretic peptides to expansions though it is not very significant. The similarity that is shared with Ph1 P 1 Timothy Grass Pollen Allergen and PNP-B is that both molecules are secreted. In terms of their relation in function very little can be deduced since their similarity is below 25%.

3.8 Conclusion

In conclusion, we observed that PNP-B shows sequence similarity to ANP, AtPNP-A, CjBAp12 (a blight-induced protein) and expansins. PNP-B shows higher sequence similarity to CjBAp12, giving it a possible pathogenesis-related response. Alignment of the active site of plant natriuretic

peptides (the region 33 - 66) with ANP showed that the function of plant natriuretic peptides is not only encoded in the primary structure. The possibility of the formation of a disulphide bond in the active site suggests that the function of PNPs is encoded in the secondary and tertiary structures.

Secondary structure comparison also showed that AtPNP-B has a highly similar secondary structure to AtPNP-A suggesting a possibility of a water homeostasis role for PNP-B in plants. There is also similarity in the predicted 3D model for AtPNP-B and AtPNP-A molecules. Wang *et al.,* (2007) and Ludidi *et al.,* (2004) have predicted a double psi β barrel (DPBB) fold for AtPNP-A using the crystal structure of the N-terminal of Ph1 P 1 Timothy Grass Pollen Allergen (accession number P43213) as a template. The DPBB fold was also predicted for AtPNP-B, suggesting a similarity in function of these two molecules. The similarity in fold of AtPNP-B with AtPNP-A, Barwin and family 45 endoglucanases infers an evolutionary relationship between these molecules.

From the bioinformatics analysis it has been shown that AtPNP-B has two possible functions; a pathogenesis-related or water homeostasis role. However, it is possible that the two roles work hand-in-hand, where AtPNP-B might work to counteract the effect of pathogens through regulation of water homeostasis. It is also possible that AtPNP-B might work directly against the pathogens that cause the infection. However it could not be concluded whether the roles work hand-in-hand or independently of each other, therefore, there is a need to investigate the actual role of AtPNP-B using purified recombinant protein. To achieve this it was necessary to develop and optimise a heterologous expression system for AtPNP-B.

CHAPTER FOUR: FUNCTIONAL CHARACTERISATION OF PNP-B IN PLANTS

4. Introduction

Preparations enriched for a specific protein are rarely easily obtained from natural host cells. Hence, recombinant protein production is frequently the sole source for purified proteins. The ribosomal machinery, located in the cytoplasm is an outstanding catalyst of recombinant protein biosynthesis (Sørensen & Mortensen, 2005). Among the systems available for heterologous protein production, the gram-negative bacterium *Escherichia coli* is one of the most attractive because of its relative simplicity, its inexpensive and fast high-density cultivation, its well characterised genetics and the large number of compatible tools available for biotechnology (Baneyx, 1999).

4.1 Essential Components of expression

A number of central elements are essential in the design of recombinant expression systems (Baneyx, 1999; Jonasson *et al.*, 2002). Expression is normally induced from a plasmid harboured by a system with compatible genetic background. The genetic elements of the expression plasmid include origin of replication (*ori*), an antibiotic resistance marker, transcriptional promoters, translation initiation regions (TIRs) as well as transcriptional and translational terminators (Sørensen & Mortensen, 2005).

4.1.1 The Replicon

The replicons of plasmids contain the origin of replication and in some cases associated *cis* acting elements (del Solar *et al.,* 1998). Most plasmid vectors used in recombinant protein expression replicate by the ColEl or p15A replicon. Plasmid copy number is controlled by the origin of replication that preferably replicates in a relaxed fashion (Baneyx, 1999). The ColEl replicon present in modern expression plasmids is derived from the pBR322 (copy number 15-20) or the pUC (copy number 500-700) family of plasmids, whereas the p15A replicon is derived from pACYC184 (copy number 10-12). These multi-copy plasmids are stably replicated and maintained under selective conditions and plasmid free daughter cells are rare (Summers, 1998).

4.1.2 Resistance Markers

The gene coding for antibiotic resistance is necessary, both for identifying transformants and to ensure antibiotic selective pressure, that is, only cells that harbour an expression vector will divide, thus preventing plasmid loss (Jonasson *et al.,* 2002). The most common drug resistance markers in recombinant expression plasmids confer resistance to ampicillin, kanamycin, chloramphenicol or tetracycline. Plasmid mediated resistance to ampicillin is accomplished by expression of β-lactamase from the *bla* gene. This enzyme is secreted to the periplasm, where it catalyses the hydrolysis of the β-lactam ring. Ampicillin present in the cultivation medium is especially susceptible to degradation, either by secreted β-lactamase, or acidic conditions in high-density cultures. The latter effect can be alleviated by the less degradation susceptible ampicillin analog, carbenicillin. Kanamycin, chloramphenicol and tetracycline interfere with protein synthesis by binding to critical areas of the ribosome. Kanamycin is inactivated in the periplasm by aminoglycoside phosphotransferases and chloramphenicol by the *cat* gene product, chloramphenicol acetyl transferase (Connell *et al.,* 2003).

4.1.3 Promoters

Table 4.1: Promoters used in high-level expression in *E. coli*.

Promoter	Induction
lac (lacUV5)	Chemical (IPTG); High temperature
trp	Chemical (IAA); High temperature
tac (hybrid)	Chemical (IPTG)
γpL	Chemical (IPTG); High temperature
T7	Chemical (IPTG); High temperature
T7-lac operator	Chemical (IPTG)
T3-lac operator	Chemical (IPTG)
T5-lac operator	Chemical (IPTG)
T4 gene 32	T4 infection
phoA	Phosphate starvation
araBAD	Chemical (L-arabinose)
xapA	Chemical (xanthosine)
cadA	Acidic pH
recA	Chemical (nalidixic acid)
lpp	Chemical (IPTG)
proU	Osmolarity
cst-1	Glucose starvation
tetA	Chemical (tetracycline)
nar	Anaerobic conditions
trc (hybrid)	Chemical (IPTG)

There are many different types of promoters (Table 4.1 above) that can affect the level of gene expression in *E. coli*. The suitability of promoters for high-level gene expression is governed by several criteria. First, the promoter must be strong, capable of protein production in excess of 10-30% of the total cellular protein. Secondly, the promoter should exhibit a minimal level of basal transcription. Basal transcription in the absence of inducer is minimized through the presence of a

suitable repressor. Minimization of basal transcription is especially important when the expression target introduces a cellular stress situation and thereby selects for plasmid loss. An early overproduction of the heterologous protein, due to a nonsilent promoter, might impair cell growth. It is therefore desirable to be able to repress the promoter during a cell growth phase to achieve high cell densities, after which the high-rate protein production would be initiated by induction of the promoter. Thirdly, promoters should be capable of induction in a simple and cost-effective manner. Promoter induction is either thermal or chemical and the most common inducer is the sugar molecule isopropyl-β-D-thiogalactopyranoside (IPTG). IPTG is an effective inducer of the powerful hybrid *tac* and *trc* promoters (Hanning & Makrides, 1998).

4.1.4 Messenger RNA

Translation initiation from the translation initiation region (TIR) of the transcribed messenger RNA requires a ribosomal binding site (RBS) including the Shine-Dalgarno (SD) sequence and a translation initiation codon (Sørensen *et al.,* 2002). The Shine-Dalgarno sequence is located 7±2 nucleotides upstream from the initiation codon, which is the canonical AUG in efficient recombinant expression systems (Ringquist *et al.,* 1992). Optimal translation initiation is obtained from mRNAs with the SD sequence UAAGGAGG. The RBS secondary structure is highly important for translation initiation and efficiency is improved by high contents of adenine and thymine (Laursen *et al.,* 2002). Translation initiation efficiency is in particular influenced by the codon following the initiation codon and adenine is abundant in highly expressed genes (Stenstrom *et al.,* 2001).

A transcription terminator placed downstream from the sequence encoding the target gene, serves as enhancement of plasmid stability by preventing transcription through the origin of replication and from irrelevant promoters located in the plasmid. Transcription terminators stabilize the

mRNA by forming a stem loop at the 3` end (Newbury *et al.*, 1987). Translation termination is preferably mediated by the stop codon UAA in *E. coli*. Increased efficiency of translation termination is achieved by insertion of consecutive stop codons or the prolonged UAAU stop codon (Poole *et al.*, 1995).

4.2 Expression Systems

A wealth of expression systems designed for various applications and compatibilities are available. Approximately 80% of the proteins used to solve three-dimensional structures submitted to the protein data bank (PDB) in 2003 were prepared in an *E. coli* expression system. The T7 pET expression system (commercialized by Novagen) is the most used in recombinant protein preparation (pET represents more that 90% of the PDB protein preparation systems). Systems using the λPL promoter/cI repressor (e.g., Invitrogen pLEX), Trc promoter (e.g., Amersham Biosciences pTrc), Tac promoter (e.g., Amersham Biosciences pGEX) and hybrid *lac*/T5 (e.g., Qiagen pQE) promoters are common (Hannig & Makrides, 1998). A radically different system is based on the *ara*BAD promoter (e.g., Invitrogen pBAD).

In this study the pQE expression system was used. High-level expression of 6xHis-tagged proteins in *E. coli* using pQE vectors is based on the T5 promoter transcription-translation system. pQE plasmids belong to the pDS family of plasmids (Bujard *et al.*, 1987) and were derived from plasmids pDS/RBSII and pDS781/RBSII-DHFRS (Stueber *et al.*, 1990). These low-copy plasmids have the following features:

- Optimized promoter-operator element consisting of phage T5 promoter (recognised by the *E. coli* RNA polymerase) and two lac operator sequences which increase lac repressor binding and ensure efficient repression of the powerful T5 promoter

- Synthetic ribosomal binding site, RBSII, for high translation rates

- 6xHis-tag coding sequence either 5` or 3` to the cloning region

- Multiple cloning site and translational stop codons in all reading frames for convenient preparation of expression construct

- Two strong transcriptional terminators: t_0 from phage lambda (Schwarz *et al.*, 1978), and T1 from *rrn*B operon of *E. coli*, to prevent read-through transcription and ensure stability of the expression construct

- β-lactamase gene (*bla*) conferring resistance to ampicillin (Sutcliffe, 1979) at 100 µg/ml. The chloramphenicol acetyl transferase (CAT) present between t_0 and T1 has no promoter and is not normally expressed

- ColE1 origin of replication (Sutcliffe, 1979).

The extremely high transcription rate initiated at the T5 promoter can only be efficiently regulated and repressed by the presence of high levels of the *lac* repressor protein. *E. coli* host strains used in the QIAexpress System use a *lac* repressor gene in *trans* or *cis* to the gene to be expressed. In the *trans* system, the host strains contain the low-copy plasmid pREP4 which confers kanamycin resistance and constitutively expresses the *lac* repressor protein encoded by the *lac* I gene (Farabaugh, 1978). The cis-repressed vectors pQE-80L, 81L, and -82L contain the *lac* I^q gene and do not require the presence of pREP4.

Expression of recombinant proteins encoded by pQE vectors is rapidly induced by the addition of IPTG which binds to the *lac* repressor protein and inactivates it. Once the *lac* repressor is inactivated, the host cell's RNA polymerase can transcribe the sequences downstream from the promoter. The transcripts produced are then translated into the recombinant protein. The special 'double operator' system in the pQE expression vectors, in combination with the high levels of the *lac* repressor protein generated by pREP4 or the $lacI^q$ gene on pQE-80L, pQE-81L, or pQE-82L, ensure tight control at the transcriptional level.

4.3 *E. coli* Host Strains

The strain or genetic background for recombinant expression is highly important. Expression strains should be deficient in the most harmful natural proteases, maintain the expression plasmid stably and confer the genetic elements relevant to the expression system (e.g., DE3). Advantageous strains for a number of individual applications are available.

In this research BL21, BL21(DE3), BL21(DE3) pLysS and JM109 were the *E. coli* host strains that were used. *E. coli* BL21 is the most common host and has proven outstanding in standard recombinant expression applications. BL21 is a robust *E. coli* B strain, able to grow vigorously in minimal media but however, non-pathogenic and unlikely to survive in host tissues and cause disease (Chart *et al.*, 2000). BL21 is deficient in ompT and lon, two proteases that may interfere with isolation of intact recombinant protein (Sørensen & Mortensen, 2005).

In the *E. coli* host strain BL21(DE3), the T7 RNA polymerase is produced from the lysogenic λ prophage DE3, and its expression is under the control of the IPTG-inducible lac UV5 promoter. The target gene cloned downstream from the T7 promoter is transcribed from the vector by

bacteriophage T7 RNA polymerase (Studier *et al.*, 1990). BL21(DE3)pLysS cells can be used with protein expression vectors that are under the control of the T7 promoter, such as pET vectors. This strain is lysogenic for λ-DE3 (Studier & Moffatt, 1986), which contains the T7 bacteriophage gene 1, encoding T7 RNA polymerase (Davanloo *et al.*, 1984) under the control of the *lac*UV5 promoter. BL21(DE3)pLysS also contains the pLysS plasmid, which carries the gene encoding T7 lysozyme. T7 lysozyme lowers the background expression level of target genes under the control of the T7 promoter but does not interfere with the level of expression achieved following induction with IPTG.

The JM109 strain is an *E. coli* K-12 derivative and is appropriate for routine cloning applications. JM109 cells are endonuclease (*endA*) deficient, greatly improving the quality of miniprep DNA, and are recombination (*recA*) deficient, improving insert stability. The *hsdR* mutation prevents the cleavage of cloned DNA by the *Eco*K endonuclease system. JM109 cells also contain the *lacI*^{q}*Z*Δ*M15* gene on the F´ episome, allowing blue-white screening for recombinant plasmids (Yanisch-Perron *et al.*, 1985).

4.4 Fusion Proteins

A wide range of protein fusion partners have been developed in order to simplify the purification and expression of recombinant proteins (Stevens, 2000). Fusion proteins or chimeric proteins usually include a partner or 'tag' linked to the passenger or target protein by a recognition site for a specific protease. Most fusion partners are exploited for specific affinity purification strategies. Fusion partners are also advantageous *in vivo*, where they might protect passengers from intracellular proteolysis (Jacquet *et al.*, 1999), enhance solubility (Davis *et al.*, 1999) or be used as specific expression reporters (Waldo *et al.*, 1999). An additional advantage of fusion proteins

is that they appear to permit the synthesis of otherwise poorly translated polypeptides. Common affinity tags are the polyhistidine tag (His-tag), which is compatible with immobilized metal-affinity chromatography (IMAC) and the glutathione S-transferase (GST) tag for purification of glutathione based resins.

In this study, the polyhistidine tag (His-tag) was used for tagging the target protein. This method utilizes immobilized metal-affinity chromatography (IMAC) to purify recombinant proteins containing a short affinity-tag consisting of polyhistidine residues. IMAC as described by Porath *et al.,* (1975) is based on the interaction between a transition metal ion (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}) immobilized on a matrix and specific amino acid side chains. Histidine is the amino acid that exhibits the strongest interaction with immobilized metal ion matrices, as electron donor groups on the histidine imidazole ring readily from coordination bonds with the immobilized transition metal. Peptides containing sequences of consecutive histidine residues are efficiently retained on IMAC. Following washing of the matrix material, peptides containing polyhistidine sequences can be easily eluted by either adjusting the pH of the column buffer or by adding free imidazole.

The method of purifying proteins with histidine residues was first described in 1987 (Hochuli *et al.,* 1987). Hochuli has developed a nitrilotriacetic acid (NTA) adsorbent for metal-chelate affinity chromatography. The NTA resin forms a quadridentate chelate and is especially suitable for metal ions with coordination numbers of six, since two valencies remain for the reversible binding of biopolymers. Dihydrofolate reductase with a poly-His-tag was successfully purified with Ni^{2+}-NTA matrices in 1988 (Hochuli *et al.,* 1988). The purification efficiency of this system was dependent on the length of the polyhistidine and the solvent system. Elution of His6-tagged proteins is effective within a range of 20-250 mM imidazole (Hefti *et al.,* 2001; Janknecht *et al.,* 1991). A disadvantage of using imidazole is that it can influence NMR experiments, competition

studies, and crystallographic trials, and the presence of imidazole often results in protein aggregates (Hefti *et al.,* 2001).

Poly-histidine affinity tags are commonly placed on either the N- or the C-terminus of recombinant proteins. Optimal placement of the tag is protein-specific. Purification using poly-histidine tags has been carried out successfully using a number of expression systems including bacteria (Chen & Hai, 1994; Rank *et al.,* 2001), yeast (Borsing *et al.*, 1997), mammalian cells (Janknecht *et al.,* 1991), and baculovirus-infected insect cells (Schmidt *et al.,* 1998).

Although the relatively small size and charge of the polyhistidine affinity tag ensures that protein activity is rarely affected, it cannot be excluded that the affinity tag may interfere with protein activity (Wu & Filutowicz, 1999). Moving the affinity tag to the opposite terminus (Halliwell *et al.,* 2001) or carrying out the purification under denaturing conditions often solves this problem.

4.5 Aim

Optimization of the heterologous expression of AtPNP-B to obtain purified recombinant protein for the functional characterisation of PNP-B in plants.

4.6 Materials & Methods

4.6.1 Amplification of the AtPNP-B gene using RT-PCR

The amplification of the AtPNP-B gene was done using the Qiagen® OneStep RT-PCR Kit (Qiagen). The *Arabidopsis thaliana* RNA samples were obtained from J. Goble (Rhodes University).

4.6.1.1 Oligonucleotide Primers

Forward and reverse oligonucleotide primers (ordered from *Inqaba Biotechnical Industries*) were designed to flank the 5' and 3' regions of the AtPNP-B coding region and to incorporate a *Bam*HI and *Sac*I restriction endonuclease site respectively for insertion of the amplified gene product into the multiple cloning site of the pQE80L expression vector (Qiagen).

Table 4.2: RT-PCR Oligonucleotide Primers. The table summarises the respective sequence, length and Tm of each primer. The incorporated *Bam*HI and *Sac*I restriction sites in the forward and reverse primers respectively are indicated in bold and underlined. The Tm values provided indicate the temperature at which the primer sequences, excluding the introduced restriction sites are expected to bind to the target sequence.

Primer	Sequence	Tm	Length
*Bam*HIF (Forward Primer)	5′- GTA **GGA TCC** ATG AGT AAA AGT ATT GTG TTT TTT TCT AC - 3'	65.3 ºC	38-mer
*Sac*IR (Reverse Primer)	5′- GTA **GAG CTC** TTA CGG AAA ATA ATC AAT AAT C- 3'	66.4 ºC	30-mer

For the RT-PCR procedure, 10 µM working solutions of the forward and reverse primers were prepared by dilution of the 100 µM stock solution with nuclease-free water (Inqaba instruction sheet supplied with primers).

4.6.1.2 Reaction Mixtures

The reaction mixtures were prepared as summarised in Table 4.3 below.

Table 4.3: RT-PCR Reaction Components and Control Mixture Compositions

Component	Volume/reaction	Negative Control	Final Concentration
RNase-free water	20.0 µl	30.0 µl	_
5X QIAGEN OneStep RT-PCR Buffer	10.0 µl	10.0 µl	1X
dNTP Mix (10 mM of each dNTP)	2.0 µl	2.0 µl	400 µM (each dNTP)
Forward Primer	3.0 µl	3.0 µl	0.6 µM
Reverse Primer	3.0 µl	3.0 µl	0.6 µM
QIAGEN OneStep RT-PCR enzyme mix	2.0 µl	2.0 µl	_
Template RNA	1.5 µg (10 µl)	_	1 pg – 2 µg/reaction
Total Volume	**50.0 µl**	**50.0 µl**	

4.6.1.3 Temperature Cycling Parameters

The PCR reactions detailed in Table 4.4 below were performed in the Eppendorf Mastercycler personal thermocycler.

Table 4.4: Thermal Cycling Conditions for RT-PCR Amplification

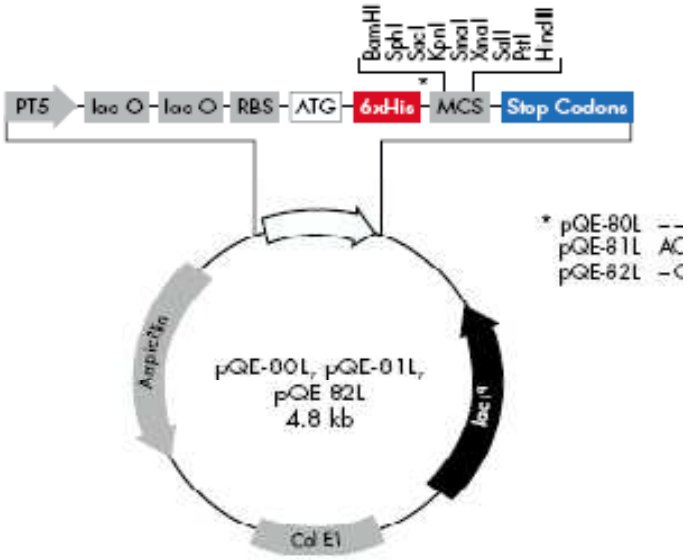
Step	Temperature	Time	Number of Cycles
Reverse Transcription	50 ºC	30 min	1
Initial PCR activation step	95 ºC	15 min	1
3-Step Cycling			
Denaturation	94 ºC	1 min	40
Annealing	55 ºC	1min	40
Extension	72 ºC	1min	40
Final Extension	72 ºC	10 min	1
Soak	4 ºC	Indefinite	1

4.6.1.4 Agarose Gel Electrophoresis of the RT-PCR Products

The RT-PCR products were resolved at 100V for 1 hour on a 1.0% agarose TAE (Tris, Acetic acid, EDTA) gel stained preceding electrophoresis by the addition of 0.5 µg/ml ethidium bromide to the gel mixture. A 10 µl sample of each reaction mixture was run with 2 µl of 6 X DNA gel loading buffer. The molecular weight marker used was 100 bp O'Gene Ruler (Fermentas). Visualization of the gel was performed in the UVP BioDoc-It™ System 2UV Transilluminator.

4.6.2 Preparation of the expression vector (pQE80L)

(a)



(b)

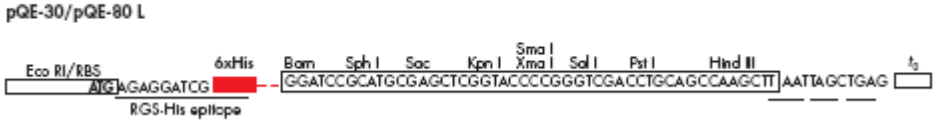


Figure 4.1(a) The map of the pQE80L vector showing the following; **PT5:** T5 promoter, **lac O**: lac operator, **RBS**: ribosome-binding site, **ATG**: start codon, **6xHis**: 6xHis tag sequence, **MCS**: multiple cloning site with restriction sites indicated, **Stop Codons**: stop codons in all three reading frames, **Col E1**: Col E1origin of replication, **Ampicillin**: ampicillin resistance gene, **lacI^q**: lacI^q repressor gene (Adapted from the QIAExpressionist). **(b)** pQE80L multiple cloning site (Adapted from the QIAExpressionist).

4.6.2.1 Transformation of the pQE80L vector into TOP10F' cells

Ten µl of TE buffer was added to 5 µg of the lyophilized pQE80L vector (Qiagen) whose map is shown in figure 4.1. The resuspended vector was transformed into chemically competent TOP10F' cells. One vial of competent TOP10F' cells (50 µl) was thawed on ice and 1 µl of the pQE80L vector was added and gently mixed by stirring with a pipette tip. The vial was incubated on ice for 30 minutes and a heat-shock treatment was applied for 30 seconds at 42 °C without shaking. The vial was immediately placed on ice and 250 µl of room temperature SOC medium

was added. The tube was tightly capped, placed on its side for better aeration and incubated for an hour at 37 °C while shaking at 200 rpm. The transformation mix (10 µl, 50 µl and 100 µl) was plated on LB plates with 100 µg/ml ampicillin (Amp) and left to grow overnight at 37 °C.

4.6.2.2 Preparation of a master plate

Ten colonies were picked from the LB plates and used in the construction of an LB Amp agar master plate. The plate was incubated at 37 ºC overnight.

4.6.2.3 Minipreparation of the pQE80L vector

The minipreparation was done using the QIAprep miniprep kit (Qiagen). Two colonies were picked from the LB Amp master plate, inoculated into 5 ml of LB medium containing 100 µg/ml ampicillin and grown at 37 ºC with shaking overnight for approximately 16 hours. The bacterial cells were harvested by centrifugation at 6800 xg for 3 min at room temperature (15 – 25 ºC) and all the supernatant was removed leaving no traces at all. The resulting pellet was resuspended in 250 µl Buffer P1, 250 µl of Buffer P2 was also added and mixed thoroughly by inverting the tube 4 – 6 times. Following addition of Buffer P2, 350 µl Buffer N3 was added and mixed immediately and thoroughly by inverting the tube 4 – 6 times. Samples were then centrifuged for 10 min at 13,000 rpm (~17,900 xg). After centrifugation, the supernatant was applied to the QIAprep spin column by decanting or pipetting, centrifuged for a minute and the flow-through was discarded. The QIAprep spin column was washed by adding 0.5 ml Buffer PB, centrifuged for a minute and the flow-through discarded. The QIAprep spin column was washed again by adding 0.75 ml Buffer PE and centrifuged for a minute. The flow-through was discarded and centrifuged for an additional 1 min to remove residual wash buffer. The QIAprep spin column was then placed in a clean 1.5 ml microcentrifuge tube. To elute DNA 50 µl of nuclease-free

water was added to the centre of each QIAprep spin column, allowed to incubate for 1 min at room temperature and then centrifuged for 1 min. From each minipreparation performed some of the eluted DNA was reserved for restriction enzyme digestion and the remaining DNA was stored at -20 ºC.

4.6.3 Construction of the pQE80L-AtPNP-B plasmid

Cloning of the AtPNP-B insert into the pQE80L expression vector was performed by double restriction enzyme digestion of the vector and insert with *Bam*HI and *Sac*I restriction enzymes as detailed below.

4.6.3.1 Double Restriction enzyme digestion of vector and insert

Table 4.5: Double Restriction enzyme digestion mixture constituents of the vector

Components	Volume/µl	Control	Control
Promega Buffer E	10 µl	10 µl	10 µl
Rnase-free water	10 µl	20 µl	20 µl
Template DNA	10 µl	_	10 µl
Promega BamHI	5 µl	5 µl	_
Promega SacI	5 µl	5 µl	_
Total	**40 µl**	**40 µl**	**40 µl**

The vector and insert were cut in separate tubes with 10U/µl *Bam*HI and *Sac*I (Promega), 10 X buffer E (Promega) in a double digest at 37 ºC for 2 hours and analyzed on a 1 % agarose gel. Restriction mixtures were prepared as detailed in Table 4.5 above and 4.6 below.

Table 4.6: Double Restriction enzyme digestion mixture constituents of the insert

Components	Volume/µl	Control	Control
Promega Buffer E	2 µl	2 µl	2 µl
Rnase-free water	14 µl	16 µl	16 µl
Template DNA	2 µl	_	2 µl
Promega BamHI	1 µl	1 µl	_
Promega SacI	1 µl	1 µl	_
Total	**20 µl**	**20 µl**	**20 µl**

4.6.3.2 Agarose gel electrophoresis of restriction enzyme digests

Restriction enzyme digests and control reactions were resolved at 100 V for 1.5 hrs on a 1 % agarose TAE gel as per the protocol detailed in section 4.6.1.4.

4.6.3.3 Gel Purification of the vector and insert

The gel purification of the vector and insert was done using the Wizard® SV Gel and PCR Clean-Up System (Promega). After the running of the gel, rapid visualisation was performed with a long-wavelength UV lamp and the two DNA fragments were excised using a sterile scalpel – one corresponding to the RT-PCR-amplified insert and the second to the vector.

The gel slices were then transferred to weighed microcentrifuge tubes and the weights recorded. The weight of the empty tube was subtracted from the total weight to obtain the weight of the gel slice. Membrane Binding Solution was then added to the gel slices at a ratio of 10 µl of solution per 10 mg of agarose gel slice. The mixture was vortexed and incubated at 50 – 65 ˚C for 10 mins until the slice was completely dissolved. For each dissolved gel slice one SV Minicolumn was placed in a Collection Tube. The dissolved gel mixture was transferred to the SV Minicolumn

assembly and incubated for 1 min at room temperature. The SV Minicolumn was then centrifuged at 16,000 xg (14,000rpm) for 1 min and the liquid in the collection tube was discarded. The column was washed by adding 700 µl of Membrane Wash Solution to the SV Minicolumn and centrifuging the column assembly for 1 min at 16,000 xg (14,000rpm). The column was washed again with 500 µl of Membrane Wash Solution and centrifuged for 5 mins at 16,000 xg. After the wash step, the collection tube was emptied and the column assembly was recentrifuged for 1 min with the microcentrifuge lid off to allow evaporation of any residual ethanol. Following centrifugation, the SV Minicolumn was transferred to a clean 1.5 ml microcentrifuge tube and 50 µl of nuclease-free water was applied directly to the centre of the column, incubated at room temperature for 1 min and then centrifuged for 1min at 16,000 xg (14,000rpm). The SV Minicolumn was discarded, some of the eluted gel purified vector and insert was reserved for ligation and the remaining eluted gel purified vector and insert was stored at -20 ˚C.

4.6.3.4 Ligation of the pQE80L vector and the AtPNP-B insert

To enable a 1:2 molar ratio of vector: insert in the ligation reaction, the reaction mixture was composed as specified in Table 4.7, below. The ligation mixture was incubated at 16 ºC for four hours followed by incubation at 4 ºC for 30 mins prior to transformation into the *E. coli* cells.

Table 4.7: Ligation Reaction of pQE80L and AtPNP-B

Reaction Constituent	Volume (µl)
BamHI/SacI-restricted pQE80L vector	4 µl
BamHI/SacI-restricted AtPNP-B (RT-PCR-amplified)	6 µl
10x Ligation Buffer (Promega)	2 µl
1.5 U/µl T4 DNA Ligase	1 µl
RNase-free water	2 µl
Total	**15** µl

4.6.3.5 Transformation of the pQE80L-AtPNP-B plasmid into TOP10F′

The pQE80L-AtPNP-B construct was transformed into chemically competent TOP10F′ cells. As described in the manual (pCR®T7 TOPO® TA Expression Kits, Invitrogen), one vial (50 µl) of the competent cells per transformation was thawed on ice and 2 µl of plasmid DNA (pQE80L-AtPNP-B) was added and gently mixed by stirring with a pipette tip. The vial was incubated on ice for 30 minutes and a heat-shock treatment was applied for 30 seconds at 42 °C without shaking. The vial was immediately placed on ice and 250 µl of room temperature SOC medium was added. The tube was tightly capped, placed on its side for better aeration and incubated for an hour at 37 °C while shaking at 200 rpm. The transformation mix was plated on LB plates with ampicillin and allowed to grow overnight at 37 °C.

4.6.3.6 Verification of the pQE80L-AtPNP-B ligation

Verification of the ligation of the pQE80L vector and the AtPNP-B insert was achieved by colony PCR followed by agarose gel electrophoresis of the colony PCR products as detailed below.

Colony PCR of transformed pQE80L-AtPNP-B

From the overnight LB Amp plates prepared in section 4.6.3.5 ten colonies were picked and used in the construction of an LB agar master plate and incubated at 37 ºC overnight. Templates for the colony PCR screening procedure were prepared through the re-suspension of the picked colonies in 10 µl of nuclease-free water. The constituents specified in Table 4.8 were added to the re-suspended cells. A PCR negative control was also included containing 10 µl of nuclease-free water in place of the template. The GoTaq® PCR Core Systems Kit (Promega) was used for the colony PCR procedure.

Table 4.8: Colony PCR reaction constituents

Component	Component Volume	Final concentration
$MgCl_2$, 25mM Solution	2.0 – 8.0 µl	1.0 – 4.0 mM
5X Colorless GoTaq Flexi Buffer OR 5X Green GoTaq Flexi Buffer	10 µl	1.0X
PCR Nucleotide Mix, 10mM each	1 µl	200 µM each
Forward primer	2 µl	0.1 – 1.0 µM
Reverse primer	2 µl	0.1 – 1.0 µM
GoTaq DNA Polymerase, 5 U/µl	0.25 µl	1.25 U/50µl
Template DNA	10 µl	<0.5 µg/50 µl
Nuclease-Free water to a final vol of	**50 µl**	

The PCR temperature cycling performed is shown in Table 4.9, below.

Table 4.9: Thermal Cycling Guidelines for Colony PCR Amplification

Step	Temperature	Time	Number of Cycles
Initial Denaturation	95ºC	2minutes	1
Denaturation	95ºC	0.5 – 1minute	25 - 35
Annealing	55ºC	0.5 – 1minute	
Extension	72 ºC	1minute/kb	
Final Extension	72 ºC	5 minutes	1
Soak	4 ºC	Indefinite	1

Agarose Gel Electrophoresis of Colony PCR Products

The colony PCR products and control reactions were resolved at 100 V for 1.5 hrs on a 1 % agarose TAE gel as per the protocol detailed in section 4.6.1.4.

4.6.3.7 Minipreparation of the pQE80L-AtPNP-B construct

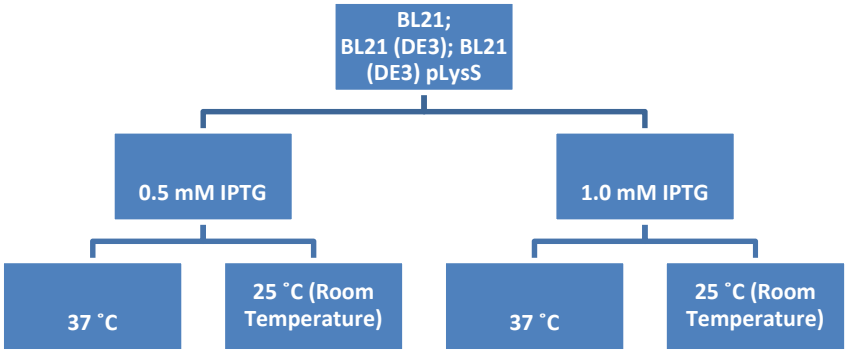
The minipreparation was done using the QIAprep miniprep kit (Qiagen). Two colonies were picked from the LB Amp master plate, prepared in section 4.6.3.6, inoculated into 5 ml of LB medium containing 100 µg/ml ampicillin and grown at 37 ºC with shaking overnight for approximately 16 hours. The bacterial cells were harvested by centrifugation at 6800 xg for 3 min at room temperature (15 – 25 ºC) and all the supernatant was removed leaving no traces at all. The resulting pellet was resuspended in 250 µl Buffer P1, 250 µl of Buffer P2 was also added and mixed thoroughly by inverting the tube 4 – 6 times. Following addition of Buffer P2, 350 µl Buffer N3 was added and mixed immediately and thoroughly by inverting the tube 4 – 6 times. Samples were then centrifuged for 10 min at 13,000 rpm (~17,900 xg). After centrifugation, the

supernatant was applied to the QIAprep spin column by decanting or pipetting, centrifuged for a minute and the flow-through was discarded. The QIAprep spin column was washed by adding 0.5 ml Buffer PB, centrifuged for a minute and the flow-through discarded. The QIAprep spin column was washed again by adding 0.75 ml Buffer PE and centrifuged for a minute. The flow-through was discarded and centrifuged for an additional 1 min to remove residual wash buffer. The QIAprep spin column was then placed in a clean 1.5 ml microcentrifuge tube. To elute DNA 50 µl of nuclease-free water was added to the centre of each QIAprep spin column, allowed to incubate for 1 min at room temperature and then centrifuged for 1 min. From each minipreparation performed some of the eluted DNA was stored at -20 ºC to be used for expression.

4.6.4 Expression Analysis of AtPNP-B

4.6.4.1 Overview of the procedure employed for optimization of expression

(a)



(b)

Figure 4.2 Overview of the conditions used during the optimization of expression

For the optimization of expression BL21, BL21 (DE3) and BL21 (DE3) pLysS were used at 0.5 mM and 1.0 mM IPTG concentrations and 37 ˚C and 25 ˚C (room temperature) as shown in the

flow diagram above (Figure 4.2a). However, using JM109 expression was done at 1.0 mM IPTG concentration and 37 ˚C only (Figure 4.2b above).

4.6.4.2 Preparation of competent *E. coli* cell cultures [BL21, BL21 (DE3), BL21 (DE3) pLysS, JM109]

A starter culture was prepared by inoculating a 2 µl aliquot of commercially prepared cells into 5ml of LB medium and grown for 4 – 6 hrs at 37 ˚C with shaking. Two millilitres of the starter culture was inoculated into 200 ml of LB medium and grown at 37 ˚C with shaking to an OD_{600} of approximately 0.4. Following this, the culture was placed on ice and the cells were harvested by centrifugation for 15 min at 2,500 xg, 4 ºC. After centrifugation, the supernatant was discarded carefully and the cell pellet re-suspended gently in 100 ml of sterile, ice cold 100 mM $CaCl_2$. The re-suspended pellet was centrifuged again for 15 min at 2,500 xg, 4 ºC and the supernatant discarded carefully. The pellet was gently re-suspended in 50 ml of sterile, ice cold 100 mM $CaCl_2$ and centrifuged for 15min at 2,500 xg, 4 ºC. Following centrifugation, the cell pellet was gently re-suspended in 5 ml of sterile, ice cold 100 mM $CaCl_2$, 20 % glycerol and 100 µl of the competent cells aliquoted into pre-chilled, sterile polypropylene tubes. The competent cells were then frozen in liquid nitrogen and immediately stored at -80 ºC.

4.6.4.3 Transformation of pQE80L- AtPNP-B for over-expression

For over-expression of AtPNP-B, the pQE80L-AtPNP-B construct was transformed into chemically competent *E. coli* cells (BL21, BL21 (DE3), BL21 (DE3) pLysS and JM109). As described in the manual (pCR®T7 TOPO® TA Expression Kits, Invitrogen). One vial (50 µl) of *E. coli* cells per transformation was thawed on ice and 2 µl of plasmid DNA (pQE80L AtPNP-B) was added and gently mixed by stirring with a pipette tip. The vial was incubated on ice for 30 minutes and a heat-shock treatment was applied for 30 seconds at 42 °C without shaking. The

vial was immediately placed on ice and 250 µl of room temperature SOC medium was added. The tube was tightly capped, placed on its side for better aeration and incubated for an hour at 37 °C while shaking at 200 rpm. The transformation mix was plated on LB plates with ampicillin and allowed to grow overnight at 37 °C.

4.6.4.4 Small scale – expression of recombinant protein

One colony that was confirmed to contain the pQE80L-AtPNP-B plasmid was picked from the master plate prepared in section 4.6.3.6 and inoculated into 2 ml of LB containing 100 µg/ml ampicillin. The culture was allowed to grow overnight at 37 ºC with shaking. The following day, 10 ml of LB media containing, 100 µg/ml Amp was inoculated with 200 µl of the overnight culture (50 x dilution). The culture was incubated at 37 °C vigorously shaking to an OD_{600} of 0.6. After reaching an OD_{600} of 0.6, a 1 ml sample was taken (T_0) and the cells harvested by centrifugation at maximum speed (13 000 rpm) for 30 seconds. The supernatant was discarded and the pellet stored at -20 °C. The remaining culture was then split into two with one culture being induced by the addition of IPTG (Isopropyl β-D-thiogalactopyranoside, Roche) to a final concentration of 0.5 mM or 1.0mM. The other culture (uninduced sample) was left incubating under the same conditions and at every hour for 5 hours a 1 ml sample was taken from both induced and uninduced cultures. Cells were harvested by centrifugation and the pellets stored at –20 °C before analysis by SDS PAGE. Glycerol stocks were prepared using the overnight culture and stored at -70 °C till needed.

4.6.5 Monitoring of Expressed Proteins

4.6.5.1 SDS-PAGE

Protein samples were analysed in a discontinuous buffer system according to the method of Laemmli (1970) on a 15 % resolving and 4 % stacking gel. The composition of resolving and stacking gels are described below:

Resolving gel: 5.0 ml 30 % Acrylamide/Bis (Bio Rad), 0.1 ml 10 % SDS (Sodium Dodecyl Sulfate, Roche), 2.5 ml 1.5 M Tris-HCl buffer (pH 8.8), 2.4 ml sterile distilled water, 50 µl 10 % APS (Ammonium persulfate, Bio Rad) and 5 µl Temed (N,N,N´N´- Tetramethylethylenediamine, Bio Rad).

Stacking gel: 1.3 ml 30 % Acrylamide/Bis (Bio Rad), 0.1 ml 10 % SDS (Roche), 2.5 ml 0.5 M Tris-HCl buffer (pH 6.8), 6.1 ml sterile distilled water, 50 µl 10 % APS (Bio Rad) and 15 µl Temed (Bio Rad).

Sample preparation for SDS-PAGE

Prior to use 50 µl of β-mercaptoethanol was added to 950 µl of sample buffer (0.5 M Tris-HCl, pH 6.8, glycerol, 10 % (w/v) SDS, 0.5 % (w/v) bromophenol blue, deionized water). The cell pellets from section 4.2.4.4 were re-suspended with 40 µl of sample buffer. The samples were then heated at 95 ºC for 5 min and then spun at 13,000 rpm in a microcentrifuge for 1 min. After spinning, 20 µl of each sample was loaded on the gel and run at 200 V for 45 min. Following the running of the gel, it was washed in 200 ml ultrapure water with gentle shaking for 15 min. The gels were stained with Imperial™ Protein Stain (Pierce) for 2 hrs and destained with 200 ml

ultrapure water overnight or until protein bands were visualized. SDS-PAGE gels to be used for western analysis were also prepared in a similar way and set aside prior to staining.

4.6.5.2 Western Blotting

Western analysis of the His-AtPNP-B was performed with the penta- and RGS- anti-His primary antibodies from mouse serum (Qiagen; Roche) and peroxidase-labeled Goat Anti-Mouse IgG secondary antibody (KPL; Roche). Resolved proteins on the SDS-PAGE gels of the induction procedure were allowed to transfer to a nitrocellulose membrane 100 V for 1 hr or 30 V overnight in a BioRad Western analysis apparatus loaded with Transfer Buffer (48 mM Tris, 39 mM glycine, 20 % (v/v) methanol, pH 9.2). The membrane was incubated in blocking buffer (10% (w/v) milk powder in TBS-T) for 1 hr at room temperature on an orbital shaker and rinsed with TBS-T. The membrane was incubated with Anti-His antibody (diluted 1: 1500 in TBS-T/1: 2500 in TBS-T) at room temperature for 1 hr on an orbital shaker. Following washing of the membrane with TBS-T the membrane was incubated with peroxidase-labeled Goat Anti-Mouse IgG secondary antibody (KPL; Roche) (diluted 1: 10 000 in TBS-T) for 1 hr at room temperature on an orbital shaker. After washing of the membrane with TBS-T, the blot was developed using the ECL Advance™ Western blotting detection kit (Amersham Biosciences) and the resulting chemiluminescence signal detected using a BioRad ChemiDoc.

4.7 Results

The *Arabidopsis thaliana* total RNA samples obtained from J. Goble (Rhodes University) were run on a 0.8% agarose gel to check on the quality of the RNA. As shown in Fig 4.3, the RNA was shown to be intact by the sharp clear bands of the 28S and 18S ribosomal RNA (rRNA) bands

that were visualized by ethidium-bromide staining. This was an indication that the RNA samples could be used for RT-PCR.

The *Arabidopsis* total RNA was quantified and the purity was also checked. Using the RNA concentration factor of 40 µg/ml, the RNA concentrations of the three RNA samples were determined as 192, 210 and 280 µg/ml. The A_{260}/A_{280} ratios of the total RNA samples indicated a negligible presence of contaminating proteins within the RNA samples.

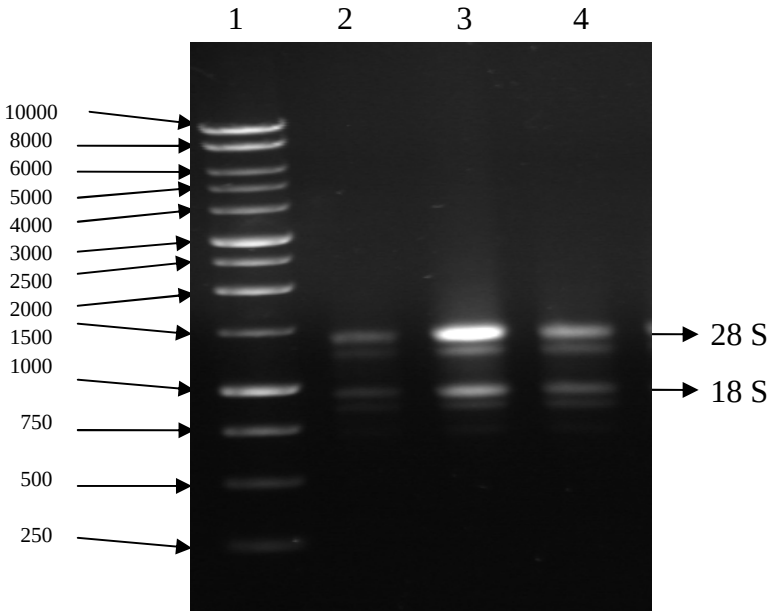


Figure 4.3 Verification of the quality of the *Arabidopsis thaliana* total RNA samples by non-denaturing 0.8% agarose gel electrophoresis. Lanes: 1 - 1kb DNA ladder; 2 – 40 mg RNA sample; 3 – 120 mg RNA sample; 4 – 80 mg RNA sample.

From the bioinformatics analysis in Chapter 3 (Bioinformatics Analysis of PNP-B) it was predicted that AtPNP-B has a 138 bp intron. The 138 bp intron was determined to still be present by the RT-PCR results which showed an RT-PCR product of approximately 510 bp (Figure 4.4 below). This is contrary to expectations and therefore a control PCR reaction was performed to

eliminate the possibility of contaminating DNA in the RNA sample. No band was detected, showing that there was no DNA contamination.

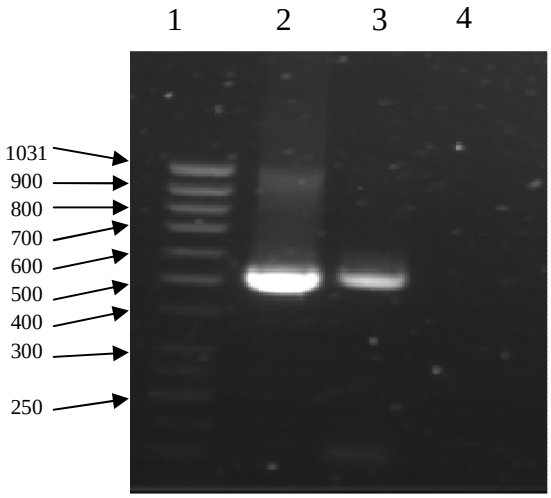


Figure 4.4 0.8% agarose gel electrophoresis of the RT-PCR amplification of AtPNP-B. Lanes: 1 – O'GeneRuler™ 50 bp DNA Ladder; 2 – RT-PCR product (2 µg RNA sample); 3 – RT-PCR product (1 µg RNA sample); 4 – Negative control (no RNA template).

Although the predicted intron was still found in the amplified product after RT-PCR, it was decided to attempt expression to determine whether a functional protein could still be expressed. For the optimization of the heterologous expression of AtPNP-B, pQE80L was was used as the expression vector. Therefore, the AtPNP-B gene was ligated into the pQE80L expression vector. Before ligation, both the RT-PCR product and the expression vector were restricted with *Bam*HI and *Sac*I restriction enzymes and run on a 1.0% agarose gel. The positions of the expression vector and the insert are shown in figure 4.5 below showing that the vector and insert were successfully restricted as they migrated to the positions corresponding to their actual sizes of approximately 4.8 kb and 510 bp respectively. The expression vector and the AtPNP-B gene were excised from the gel for gel purification. The gel purified vector and insert were ligated and

transformed into competent TOP10F` cells. Ten colonies resulting from the transformation procedure were screened for the pQE80L-AtPNP-B construct by colony PCR. The colony PCR products were analysed by 1.0% agarose electrophoresis, as presented in figure 4.6 below.

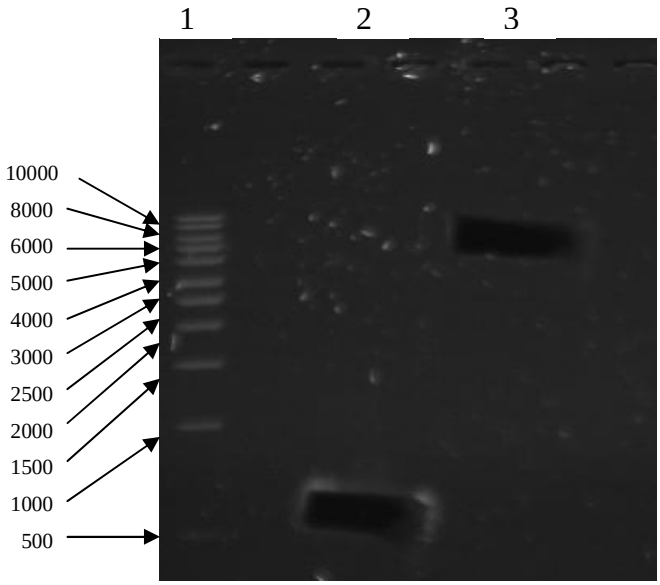


Figure 4.5 1.0% agarose gel electrophoresis of the doubly restricted (BamHI and SacI) pQE80L and the AtPNP-B gene. Lanes: 1 – Perfect DNA™ 1kb Ladder; 2 – Position of the excised doubly restricted RT-PCR amplified AtPNP-B gene; 3 – Position of the excised doubly restricted pQE80L expression vector.

Following colony PCR, the AtPNP-B inserted into the expression vector was expected to be amplified and visualised on an agarose gel at a size that corresponds to approximately 510 bp as depicted in figure 4.6 below. All the colonies showed the presence of the pQE80L-AtPNP-B construct.

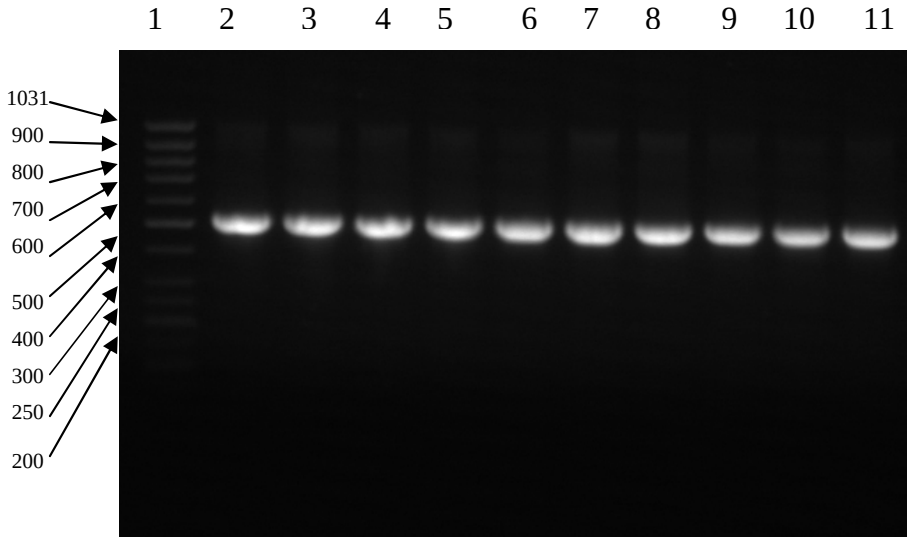
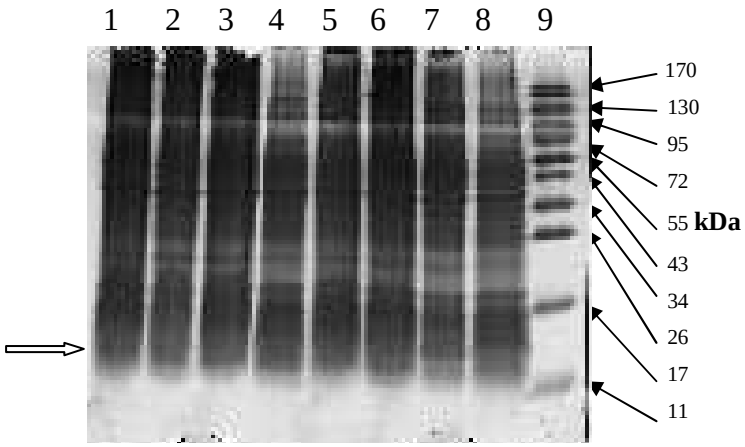


Figure 4.6 Verification of the pQE80L-AtPNP-B ligation by colony PCR, as visualised by 1.0% agarose electrophoresis. Lanes: 1 – O'GeneRuler™ 50 bp DNA Ladder; Colony PCR products from; 2 – colony 1; 3 – colony 2; 4 – colony 3; 5 – colony 4; 6- colony 5; 7 – colony 6; 8 – colony 7; 9 – colony 8; 10 – colony 9; 11 – colony 10.

After the confirmation of the insert in the expression vector, only one of the ten colonies was picked for the expression of AtPNP-B. The expression of AtPNP-B encoded by the pQE80L-AtPNP-B construct was done using the conditions shown in figure 4.2 under IPTG induction. The *E. coli* hosts used were BL21, BL21 (DE3), BL21 (DE3) pLysS and JM109. Protein fractions were collected from the induced and non-induced samples. When SDS-PAGE analysis was performed on all the SDS-PAGE gels there was no induced protein that was observed and typical gels showing the SDS-PAGE results is presented in figure 4.7 below. The assumption was that the protein had been induced in very small quantities that could not be resolved on the SDS-PAGE gels, thus western blot analysis was also carried out on the gels. However, the western analysis also yielded no results and typical western blot pictures are presented in figure 4.8 below.

(a)



(b)

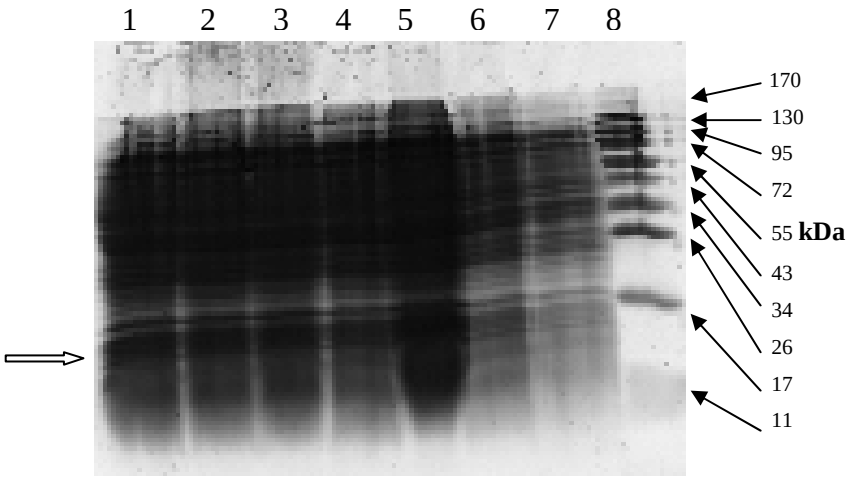
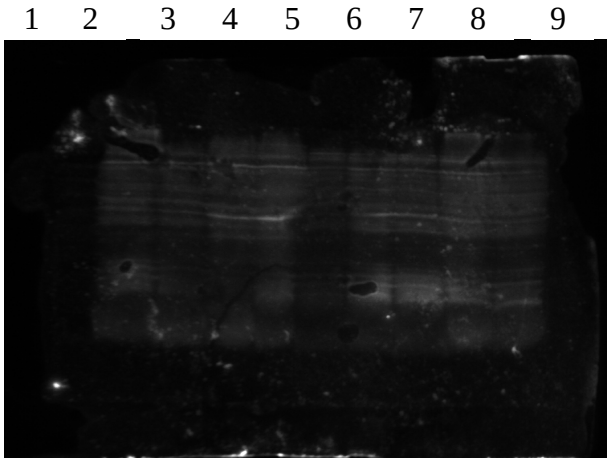


Figure 4.7 SDS-PAGE of protein fractions.

(a) IPTG-induced *E. coli* BL21(DE3) cells; Lanes: 1 – T_4 Induced; 2 – T_3 Uninduced; 3 – T_3 Induced; 4 – T_2 Uninduced; 5 – T_2 Induced; 6 – T_1 Uninduced; 7 – T_1 Induced; 8 – T_0; 9 – peqGOLD Protein Marker IV.

(b) IPTG-induced JM109 cells; Lanes: 1 - T_3 Uninduced; 2 - T_3 Induced; 3 - T_2 Uninduced; 4 - T_2 Induced; 5 - T_1 Uninduced; 6 - T_1 Induced; 7 - T_0; 8 - peqGOLD Protein Marker IV. The unfilled block arrow indicates the expected position of the induced 13kDa His-tagged AtPNP-B protein.

(a)



(b)

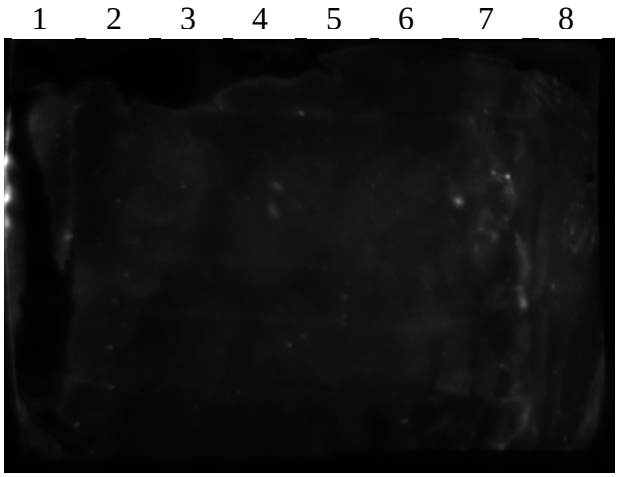


Figure 4.8 Western analysis of resolved proteins.

(a) IPTG-induced *E. coli* BL21 cells; Lanes: 1 – T_4 Induced; 2 – T_3 Uninduced; 3 – T_3 Induced; 4 – T_2 Uninduced; 5 – T_2 Induced; 6 – T_1 Uninduced; 7 – T_1 Induced; 8 – T_0; 9 – peqGOLD Protein Marker IV.

(b) IPTG-induced JM109 cells; Lanes: 1 - T_3 Uninduced; 2 - T_3 Induced; 3 - T_2 Uninduced; 4 - T_2 Induced; 5 - T_1 Uninduced; 6 - T_1 Induced; 7 - T_0; 8 - peqGOLD Protein Marker IV.

4.8 Discussion

The presence of the predicted 138 bp intron of AtPNP-B was confirmed by the RT-PCR results giving an RT-PCR product of approximately 510 bp. This could mean that AtPNP-B is a pseudogene or the predicted intron is not correct and the gene could express a functional protein. Expression studies under different conditions was therefore attempted to determine if a functional protein could be expressed.

For expression, the T5-based pQE80L expression vector was used. The T5 bacteriophage promoter has the advantage of being expressed by bacterial RNA polymerase, unlike the T7 RNA polymerase (Zomorrodipour *et al.,* 2004). *E. coli* strains used in the QIAexpress System use a lac repressor gene in *trans* or *cis* to the gene to be expressed. In the *trans* system, the host strains contain the low-copy plasmid pREP4 which confers kanamycin resistance and constitutively expresses the lac repressor protein encoded by the lac I gene (Farabaugh, 1978). Lac I is a weakly expressed gene and a 10-fold enhancement of the repression is achieved when the over-expressing promoter mutant $lacI^q$ is employed (Calos, 1978). Therefore, the cis-repressed vector, pQE80L which contains the $lacI^q$ gene and does not require the presence of pREP4 was used in this study. The special "double operator" system in pQE80L, in combination with the high levels of the lac repressor protein generated by the $lacI^q$ gene ensures tight control at the transcriptional level thus suppressing basal protein expression completely before induction.

The other advantage of a T5-based expression vector is that it is recognized by an endogenous *E. coli* RNA polymerase (Tyler *et al.,* 2005) while the T7 promoter is recognized by a highly specific viral T7 RNA polymerase (Dubendorff & Studier, 1991). The cis-repressed, pQE80L series of vectors can also easily be used with any *E. coli* host strain. Thus, the pQE80L

expression vector was used for expression in the following *E. coli* strains; BL21, BL21(DE3), BL21(DE3) pLysS and JM109.

BL21 and its derivatives, BL21(DE3) and BL21(DE3) pLysS were used in this research for expression at the conditions shown in figure 4.2. BL21 and its derivatives, BL21(DE3) and BL21(DE3) pLysS lack the OmpT and lon, two proteases that may interfere with isolation of intact recombinant protein. The lon protease is an intracellular protease that *E. coli* makes to degrade abnormal proteins and the ompT protease is a protease that *E. coli* makes to degrade extracellular proteins. Cells deficient in these proteases accumulate recombinant proteins at a high rate and are less likely to degrade some proteins during purification (Moffatt & Studier, 1987). BL21(DE3) pLysS is also less leaky, but the strain tends to produce less recombinant protein after induction (Studier *et al.*, 1990). However, BL21 and its derivatives were not successful in inducing the expression of AtPNP-B. It was concluded that the level of recombinant protein was too low to be detected on the SDS-PAGE gel or the mRNA was not compatible with the bacterial expression machinery of BL21 and its derivatives. Thus, western analysis was also done on the SDS-PAGE gels to check if there was any induction at all. The westerns also yielded negative results.

Since BL21 cells have the T7 promoter expression system and pQE80L is a T5-based expression vector it could have contributed to the expression not being successful. Thus, JM109, a T5-based *E. coli* strain was also used for the heterologous expression of AtPNP-B. JM109 also contains the $lacI^q$ gene (Yanisch-Perron *et al.*, 1985) which is compatible with the mutant lacI gene in the pQE80L expression vector. The use of JM109 was expected to give an expression product but no induction was observed from the SDS-PAGE gels. Western analysis was also done on the SDS-PAGE gels. However, no induction was observed as well.

Introns interrupt the continuity of many eukaryotic genes and therefore their removal by splicing is a crucial step in gene expression. Since the amplified AtPNP-B gene contained an intron it was also crucial to splice the intron out. The presence of the intron in the AtPNP-B gene could have contributed to the expression of the protein being unsuccessful because the *E. coli* host strains do not have splicing systems to excise the intron from the transcript resulting in the gene not being expressed.

Premature translational termination could have been another factor that resulted in the expression being unsuccessful. The presence of sequences in the AtPNP-B gene that act as termination signals in *E. coli* could also have resulted in premature translational termination. These sequences will be perfect in the plant cell but in the bacteria will result in premature termination and loss of gene expression.

4.9 Conclusion

Since expression was unsuccessful for the PNP-B gene obtained from *Arabidopsis thaliana* it was considered important to check in other plants whether the PNP-B gene is present and whether there were any plants that had an intronless PNP-B gene. If such a plant could be found then it would be used for the heterologous expression of PNP-B for functional characterisation.

It was also realised that there was a need to carry out an evolutionary study of the gene if it was going to be found in any other plants. In line with other studies of the Plant Stress group at UFH which focus on cereal studies, it was decided to evaluate the presence of PNP-B genes in cereals and compare these results with other dicotyledons.

CHAPTER FIVE: EVOLUTIONARY RELATIONSHIP OF PNP-B IN PLANTS

5. Introduction

The process of evolution changes the distribution of genotypes and phenotypes in successive generations (Lesk, 2005). A genotype is an organism's genetic information. All other observable features of an organism - macroscopic and biochemical – comprise its phenotype (Griffiths *et al.*, 2008). The genotype is inherited from a parent or parents, subject to modification by mutation or lateral transfer of genetic material (Lesk, 2005).

Evolution alters the composition and distribution of the gene pools and phenotypes in populations. Natural selection (enhanced reproduction by 'fitter' individuals) is the most important mechanism of evolution (Stearns & Hoekstra, 2005). Another is genetic drift, the random change in allelic frequencies, not in response to selection. Mechanisms that produce genetic variety such as mutations, recombinations, gene duplication, followed by divergence and gene flow create the potential for evolution (Alexandrov, 1994; Lesk, 2005).

Microevolution refers to relatively small changes in a few genes, leading in most cases to relatively small changes in phenotypes. Microevolution affects the individuals within a population. On the other hand, macroevolution refers to larger-scale changes in populations as a whole, including formation of new species (Lesk, 2005).

The information archive in each organism is the genetic material, DNA, or, in some viruses, RNA. Implicit in the structure of the DNA are mechanisms for self-replication and for translation

of genes into proteins. However, not all DNA is expressed as proteins or structural RNA. Most genes in higher organisms contain internal untranslated regions, or introns. Proteins are the molecules responsible for much of the structure and activities of organisms. Near-perfect replication is essential for stability of inheritance; but some imperfect replication, or mechanism for import of foreign genetic material, is also essential for evolution to take place (Lesk, 2005).

5.1 Aim

To investigate the evolutionary development of PNP-B in plants and sequence the coding region of the peptide in different evolutionary-related plants.

5.2 Materials & Methods

5.2.1 Plants Used in This Study

The following plants were used in this study;

Wheat (Betta), Wheat (Tugella), Barley (Clipper), Maize (CRN 3505), Cabbage (Cape Spitz/Sugar loaf), Tobacco, Millet, Sorghum, Sugar beans, Tomato (Tamatie-Floradade, PEZ8732), Mustard (Brassica nigra), Yellow maize (Sahara, PEZ8716) and Maize (Early Pearl) were grown in the glass house under natural seasonal daylight. The plants were harvested after two weeks. Moss and cycad were harvested from already existing plants.

5.2.2 Isolation of Total RNA from the Plants

Using the SV Total Isolation System (Promega), RNA was isolated from the two-week old plants, fern and cycad. All instruments and equipment used in the isolation procedure were treated with RNAse Away (Molecular BioProducts) and/or autoclaved, followed by rinsing with DEPC-treated dddH_2O.

5.2.2.1 Preparation of the plant tissue

The leaves from the plants were frozen in liquid nitrogen and ground vigorously in a pestle and mortar to a fine, dry powder. Two reaction tubes for each plant were comprised as detailed in Table 5.1 below and treated with RNA Lysis Buffer and RNA Dilution Buffer (Promega) as specified.

Table 5.1: Preparation of plant tissue for RNA isolation

Sample	Plant Tissue (mg)	RNA Lysis Buffer (µl)	RNA Dilution Buffer (µl)
1	60	175	350
2	120	175	350

Samples were centrifuged for 10 minutes at approximately 12,000 xg (14,000 rpm) and the resulting lysate transferred to a sterile microcentrifuge tube.

5.2.2.2 RNA Isolation Procedure

After the addition of 200 µl of 95% ethanol to the resulting lysate, the mixture was transferred to a Spin Column assembly and centrifuged at approximately 12,000 xg for 1 minute. The resulting flow-through solution was discarded and 600 µl of SV RNA Wash Solution (Promega) was added to the column, followed by an additional 1 minute of centrifugation at 12,000 xg. The

flow-through was discarded and 50 µl of DNAse incubation mix, comprised of 40 µl of Yellow Core Buffer, 5 µl of 0.09 M $MnCl_2$ and 5 µl of DNAse I (Promega) was added to the column. The column was centrifuged for 1 minute as before and the flow-through was discarded. The column was washed with 600 µl of SV RNA Wash Solution as detailed before. An additional wash was performed with 200 µl of SV RNA Wash Solution, followed by 2 minutes of centrifugation at 12,000 xg. Elution of the RNA into a sterile microcentrifuge was done through the addition of 50 µl of nuclease-free dddH_2O (Promega) to the column and centrifugation at 12,000 xg for 1 minute following 1 minute of incubation at room temperature. A 20 µl sample from, each isolation performed was retained for quantification and agarose gel electrophoresis and the remaining isolated RNA was stored at -20˚C.

5.2.2.3 Quantification of isolated RNA

The quality and quantity of isolated RNA was simultaneously assessed through a measurement of the A_{260} and A_{280} of each sample obtained from the isolation procedure. A 50x dilution of each sample was prepared in dddH_2O and the absorbance read using the BioMate3 Thermo Spectronic spectrophotometer. dddH_2O was used as the blank.

5.2.2.4 Agarose gel electrophoresis of the isolated RNA

Agarose gel electrophoresis was performed for verification of the isolation of RNA from the plants as well as to assess the integrity of the RNA. The RNA was resolved at 100 V for 1 hour on a 1% agarose TAE gel containing ethidium bromide to a final concentration of 0.5 µg/ml. A 10 µl sample from each isolation was run with 2 µlof 6x DNA gel loading buffer (30% glycerol, 0.25% bromophenol blue). Visualization was done using the UVP BioDoc-It™ System 2UV Transilluminator.

5.2.3 Amplification of the PNP-B gene by RT-PCR

The total RNA extracted from the plants was used in a 50µl RT-PCR reaction volume to amplify the PNP-B gene. The RT-PCR reaction was carried out using a RT-PCR Kit (Qiagen) according to the manufacturer's instructions.

RT-PCR was optimised with respect to RNA quantities and temperature. The RT-PCR reactions were optimised at 1.0, 1.5 and 2.0 µg of RNA and 50 ˚C, 55 ˚C and 60 ˚C. The optimal conditions for amplification of the PNP-B gene were noted to be at 1.5 µg RNA concentration and 55 ˚C (as shown in figure 5.3 and 5.4 below).

5.2.3.1 Oligonucleotide Primers

The forward and reverse oligonucleotide primers (ordered from Inqaba Biotechnical Industries) used to amplify the PNP-B coding region, were designed based on the *Arabidopsis thaliana,* AtPNP-B gene sequence.

Table 5.2: RT-PCR Oligonucleotide Primers. The table summarises the respective sequence, length and Tm of each primer. The incorporated *Bam*HI and *Sac*I restriction sites in the forward and reverse primers respectively are indicated in bold and underlined. The Tm values provided indicate the temperature at which the primer sequences, excluding the introduced restriction sites are expected to bind to the target sequence.

Primer	Sequence	Tm	Length
*Bam*HIF (Forward Primer)	5′- GTA **GGA TCC** ATG AGT AAA AGT ATT GTG TTT TTT TCT AC - 3'	65.3 ºC	38-mer
*Sac*IR (Reverse Primer)	5′- GTA **GAG CTC** TTA CGG AAA ATA ATC AAT AAT C- 3'	66.4 ºC	30-mer

For the RT-PCR procedure, 10 µM working solutions of the forward and reverse primers were prepared by dilution of the 100 µM stock solution with nuclease-free water (Promega).

5.2.3.2 Reaction Mixtures

The reaction mixtures were prepared comprising of the constituents summarised in Table 5.3 below.

Table 5.3: RT-PCR Reaction Components and Control Mixture Compositions

Component	Volume/reaction	Negative Control	Final Concentration
RNase-free water	20.0 µl	30.0 µl	_
5X QIAGEN OneStep RT-PCR Buffer	10.0 µl	10.0 µl	1X
dNTP Mix (10 mM of each dNTP)	2.0 µl	2.0 µl	400 µM (each dNTP)
Forward Primer	3.0 µl	3.0 µl	0.6 µM
Reverse Primer	3.0 µl	3.0 µl	0.6 µM
QIAGEN OneStep RT-PCR enzyme mix	2.0 µl	2.0 µl	_
Template RNA	1.5 µg (10 µl)	_	1 pg – 2 µg/reaction
Total Volume	**50.0 µl**	**50.0 µl**	

5.3.2.3 Temperature Cycling Parameters

The PCR reactions detailed in Table 5.4 were performed in the Eppendorf Mastercycler personal thermocycler.

Table 5.4: Thermal Cycling Conditions for RT-PCR Amplification

Step	Temperature	Time	Number of Cycles
Reverse Transcription	50 ºC	30 min	1
Initial PCR activation step	95 ºC	15 min	1
3-Step Cycling			
Denaturation	94 ºC	1 min	40
Annealing	55 ºC	1min	40
Extension	72 ºC	1min	40
Final Extension	72 ºC	10 min	1
Soak	4 ºC	Indefinite	1

5.2.3.4 Agarose Gel Electrophoresis of the RT-PCR Products

The RT-PCR products were resolved at 100V for 1 hour on a 1.0% agarose TAE (Tris, Acetic acid, EDTA) gel stained preceding electrophoresis by the addition of 0.5 µg/ml ethidium bromide to the gel mixture. A 10 µl sample of each reaction mixture was run with 2 µl of 6 X DNA gel loading buffer. The molecular weight marker used was 100 bp O'Gene Ruler (Fermentas). Visualization of the gel was performed in the UVP BioDoc-It™ System 2UV Transilluminator.

5.2.4 Construction of the pDrive-PNP-B plasmid

The RT-PCR isolated cDNA fragments were ligated into a pDrive cloning vector (map is shown in figure 5.1) using the Qiagen® PCR Cloning Kit as per manufacturer's instructions.

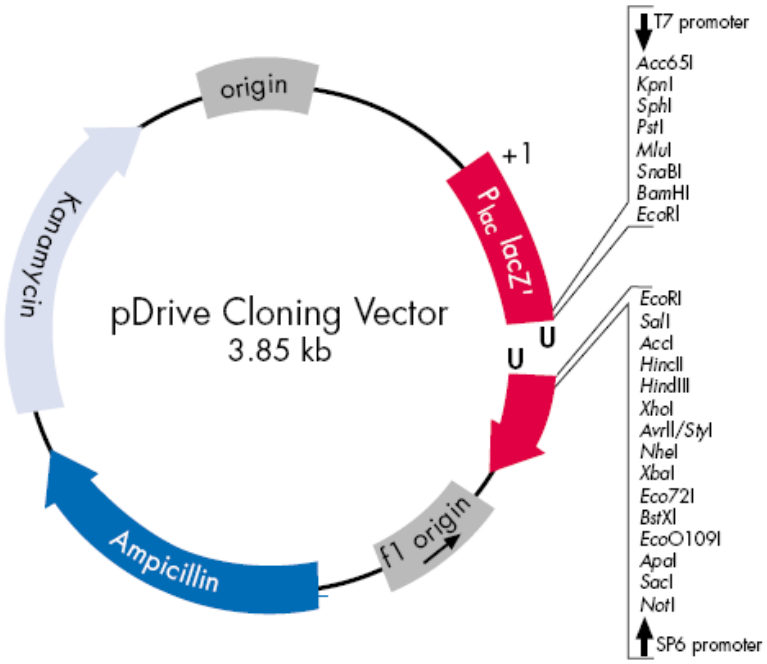


Figure 5.1 Representation of the linearized pDrive Cloning Vector with U overhangs. The unique restriction endonuclease recognition sites on either side of the cloning site are listed. (http://www1.qiagen.com/literature/pDrive/pcr_cloning21.pdf)

5.2.4.1 Ligation of the pDrive cloning vector and the PNP-B insert

To enable a 1:2 molar ratio of vector: insert in the ligation reaction, the reaction mixture was comprised as specified in Table 5.5, below. The ligation mixture was incubated at 16 ºC for four hours followed by incubation at 4 ºC for 30mins prior to transformation into competent TOP10F` cells (Invitrogen).

Table 5.5: Ligation Reaction of the pDrive cloning vector and PNP-B

Component	Volume/Reaction
pDrive Cloning Vector (50 ng/µl)	1 µl
PCR Product	2 µl
Distilled Water	2 µl
Ligation Master Mix, 2X	5 µl
Total	**10 µl**

5.2.4.2 Transformation of the pDrive-PNP-B plasmid into TOP10F′

The pDrive-PNP-B construct was transformed into chemically competent TOP10F′ cells. As described in the manual (pCR®T7 TOPO® TA Expression Kits, Invitrogen), one shot vial of the competent cells (50 µl) per transformation was thawed on ice and 2 µl of plasmid DNA (pDrive-PNP-B) was added and gently mixed by stirring with a pipette tip. The vial was incubated on ice for 30 minutes and a heat-shock treatment was applied for 30 seconds at 42 °C without shaking. The vial was immediately placed on ice and 250 µl of room temperature SOC medium was added. The tube was tightly capped, placed by its side for better aeration and incubated for an hour at 37 °C while shaking at 200 rpm. The transformation mix was plated on LB plates with ampicillin (100 µg/ml) and allowed to grow overnight at 37 °C.

5.2.4.3 Verification of the pDrive-PNP-B ligation

Verification of the ligation of the pDrive cloning vector and the PNP-B insert was achieved by colony PCR followed by agarose gel electrophoresis of the Colony PCR products as detailed below.

Colony PCR of transformed pQE80L-AtPNP-B

From the overnight LB Amp plates prepared in section 5.2.4.2 above ten colonies were picked and used in the construction of an LB agar master plate containing 100 µg/ml ampicillin and incubated at 37 ºC overnight. Templates for the colony PCR screening procedure were prepared through the re-suspension of the picked colonies in 10 µl of nuclease-free water. The constituents specified in Table 5.6 were added to the re-suspended cells. A PCR negative control was also similarly comprised containing 10 µl of nuclease-free water in place of the template. The GoTaq® PCR Core Systems Kit (Promega) was used for the colony PCR procedure.

Table 5.6: Colony PCR reaction constituents

Component	Component Volume	Final concentration
$MgCl_2$, 25mM Solution	2.0 – 8.0 µl	1.0 – 4.0 mM
5X Colorless GoTaq Flexi Buffer	10 µl	1.0X
OR 5X Green GoTaq Flexi Buffer		
PCR Nucleotide Mix, 10mM each	1 µl	200 µM each
Forward primer	2 µl	0.1 – 1.0 µM
Reverse primer	2 µl	0.1 – 1.0 µM
GoTaq DNA Polymerase, 5 U/µl	0.25 µl	1.25 U/50 µl
Template DNA	10 µl	<0.5 µg/50 µl
Nuclease-Free water to a final vol of	**50 µl**	

The PCR temperature cycling performed is shown in Table 5.7, below.

Table 5.7: Thermal Cycling Guidelines for Colony PCR Amplification

Step	Temperature	Time	Number of Cycles
Initial Denaturation	95ºC	2minutes	1
Denaturation	95ºC	0.5 – 1minute	25 - 35
Annealing	55ºC	0.5 – 1minute	
Extension	72 ºC	1minute/kb	
Final Extension	72 ºC	5 minutes	1
Soak	4 ºC	Indefinite	1

Agarose Gel Electrophoresis of Colony PCR Products

The colony PCR products and control reactions were resolved at 100 V for 1.5 hrs on a 1 % agarose TAE gel as per the protocol detailed in section 5.2.3.4.

5.2.4.4 Minipreparation of the pDrive-PNP-B construct

The minipreparation was done using the QIAprep miniprep kit (Qiagen). Two colonies were picked from the LB Amp master plate, prepared in section 5.2.4.2 above, inoculated into 5 ml of LB medium containing 100 µg/ml ampicillin and grown at 37 ºC with shaking overnight for approximately 16 hours. The bacterial cells were harvested by centrifugation at 6800 xg for 3 min at room temperature (15 – 25 ºC) and all the supernatant was removed leaving no traces at all. The resulting pellet was resuspended in 250 µl Buffer P1, 250 µl of Buffer P2 was also added and mixed thoroughly by inverting the tube 4 – 6 times. Following addition of Buffer P2, 350 µl Buffer N3 was added and mixed immediately and thoroughly by inverting the tube 4 – 6 times. Samples were then centrifuged for 10 min at 13,000 rpm (~17,900 xg). After centrifugation, the

supernatant was applied to the QIAprep spin column by decanting or pipetting, centrifuged for a minute and the flow-through was discarded. The QIAprep spin column was washed by adding 0.5 ml Buffer PB, centrifuged for a minute and the flow-through discarded. The QIAprep spin column was washed again by adding 0.75 ml Buffer PE and centrifuged for a minute. The flow-through was discarded and centrifuged for an additional 1 min to remove residual wash buffer. The QIAprep spin column was then placed in a clean 1.5 ml microcentrifuge tube. To elute DNA 50 µl of nuclease-free water was added to the centre of each QIAprep spin column, allowed to incubate for 1 min at room temperature and then centrifuged for 1 min. From each minipreparation performed some of the eluted DNA was stored at -20 ºC to be used for sequencing and the rest was used for quantifying the DNA.

5.2.5 Sequencing of the PNP-B gene

Sequencing was done using the ABI Prism™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems) at the Central Analytical Facility, University of Stellenbosch. The resulting sequence chromatograms were analysed and all anomalies were corrected before verification of the nucleotide sequences using BLAST (blastn; Search nucleotide database using a nucleotide sequence).

5.2.6 Search for other sequences similar to PNP-B

BLAST (tblastn; Search translated nucleotide database using a protein query and tblastx; Search translated nucleotide database using a translated nucleotide query) was used to get other sequences that are similar to the AtPNP-B sequence from other plant genomes on "The Gene Index Project" (http://compbio.dfci.harvard.edu/tgi/). The default settings were used in the BLAST searches. The nucleotide sequences were then translated into all 6 reading frames using

the GeneRunner Software ver 3.05 (Hasting Software, Inc). The applicable protein sequence was selected, copied to a text editor and saved in a FASTA format.

5.2.7 Sequence Alignments and Evolutionary Trees

Sequence alignments were done using CLUSTALW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) (Larkin *et al*., 2007) using the relevant text files that were obtained in section 5.2.6 above for inputting the sequences. Phylogenetic trees were constructed using MEGA4 (Tamura *et al.*, 2007). Default settings were used in all the alignments and the construction of trees.

5.2.8 Pattern Searches

The signature pattern was designed from the alignment of AtPNP-B, OsJPNP-B, CjBAp12 and WhBPNP-B. Motif searches were performed against the Prosite (http://au.expasy.org/prosite/) database as described by Sigrist *et al.*, (2002). Default settings were used for all the pattern searches.

5.3 Results

The isolated total RNA samples from all the plant samples were run on a 0.8% agarose gel to check on the quality of the RNA. The RNA was shown to be intact by the presence of sharp clear bands representing the 28S and 18S ribosomal RNA (rRNA) bands that were visualized by ethidium-bromide staining (Figure 5.2). This was an indication that the RNA samples could be used for RT-PCR.

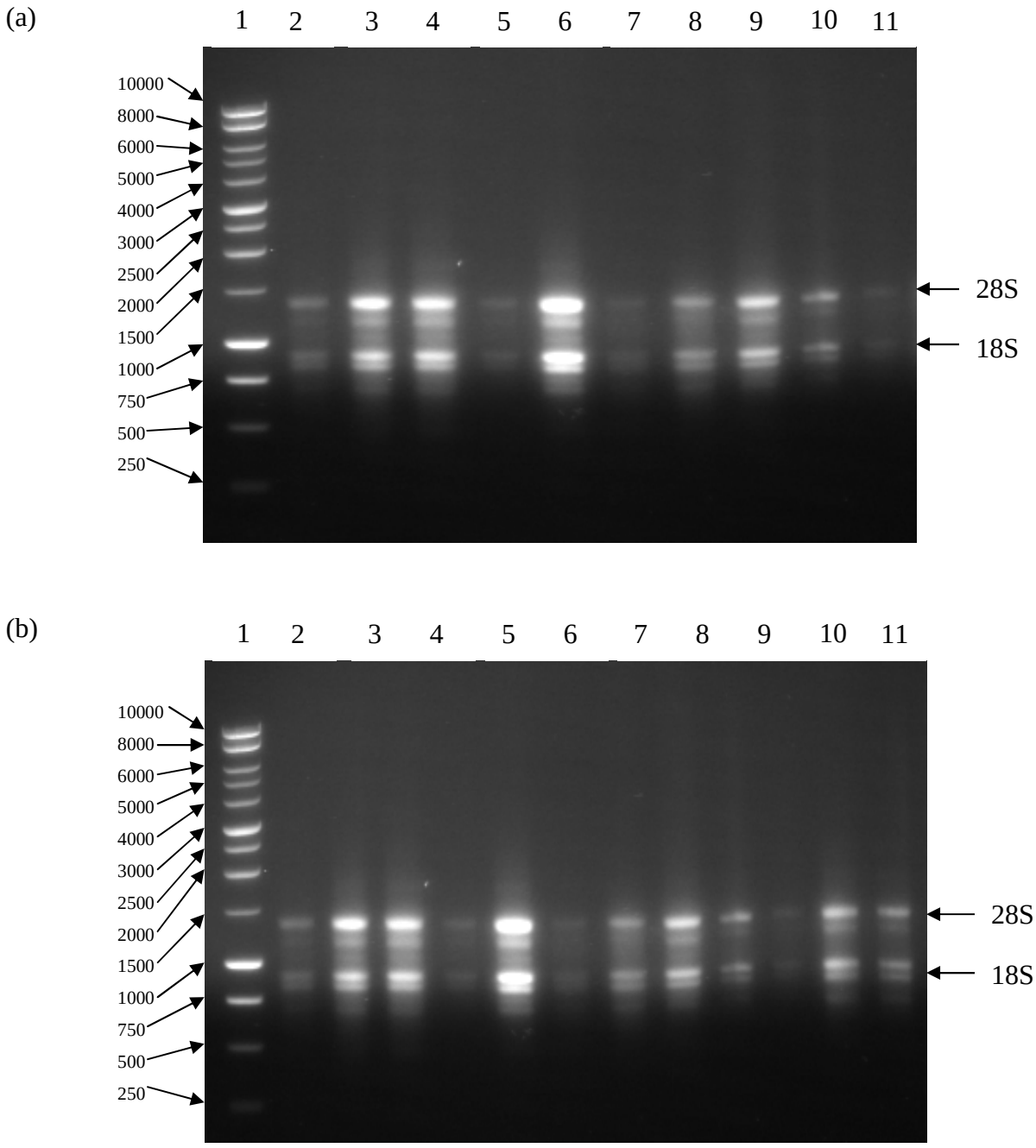


Figure 5.2 Verification of the quality of the isolated total RNA from the plants by non-denaturing 0.8% agarose gel electrophoresis.

(a) Lanes: 1 - 1kb DNA ladder; 2 - 60 mg wheat (betta) RNA sample; 3 – 120 mg wheat (betta) RNA sample; 4 – 120 mg wheat (tugella) RNA sample; 5 – 60 mg wheat (tugella) RNA sample; 6 – 60 mg maize (early pearl) RNA sample; 7 – 120 mg maize (early pearl) RNA sample; 8 – 60 mg moss RNA sample; 9 – 120 mg moss RNA sample; 10 – 120 mg cycad RNA sample; 11 – 60 mg cycad RNA sample.

(b) Lanes: 1 – 1kb DNA ladder; 2 – 60 mg barley (clipper) RNA sample; 3 – 120 mg barley (clipper) RNA sample; 4 – 120 mg cabbage RNA sample; 5 – 60 mg cabbage RNA sample; 6 – 120 mg tobacco RNA sample; 7 – 60 mg tobacco RNA sample; 8 – 60 mg sorghum RNA sample; 9 – 120 mg sorghum RNA sample; 10 – 120 mg tomato RNA sample; 11 – 60 mg tomato RNA sample; 12 – 60 mg mustard RNA sample; 13 – 120 mg mustard RNA sample.

The isolated total RNA was quantified and the purity was also checked. Using the RNA concentration factor of 40 µg/ml, the RNA concentrations of the isolated RNA samples were determined as shown in table 5.8 below. The concentrations of RNA from plants where the PNP-B gene could not be amplified are not shown in the table. The A_{260}/A_{280} ratios of the isolated total RNA samples indicated a negligible presence of contaminating proteins within the RNA samples.

Table 5.8: Quantification of the isolated total RNA from the plants

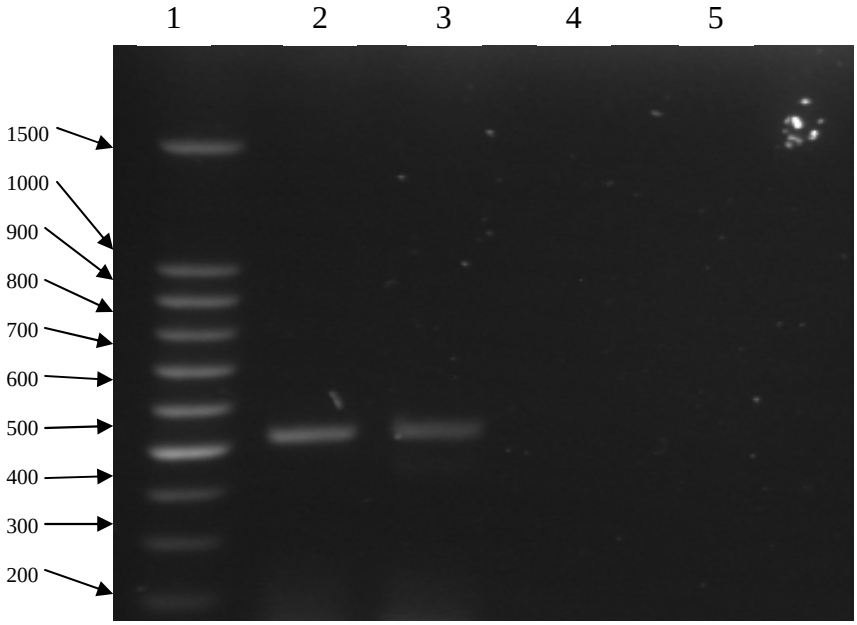
Sample	Tissue mass (mg)	A_{260}	A_{280}	Ratio	Conc. (µg/ml)
Maize (Early Pearl)	60	0.126	0.059	2.136	116.0
	120	0.103	0.049	2.102	94.29
Wheat (Betta)	60	0.169	0.080	2.112	155.0
	100	0.411	0.201	2.045	372.3
Wheat (Tugella)	60	0.179	0.084	2.131	164.7
	100	0.334	0.162	2.062	303.5
Moss	60	0.044	0.028	1.571	35.2
	120	0.083	0.048	1.729	66.40
Cycad	60	0.035	0.021	1.667	28.00
	120	0.071	0.042	1.700	56.80

From the plants that were used in this study, the presence of the PNP-B gene was noted from wheat (betta), wheat (tugella) and maize (early pearl). The 138 bp intron was shown to be present from the RT-PCR results which showed an RT-PCR product of approximately 510 bp (Figure 5.3

below). However, no PNP-B product could be amplified from RNA isolated from the plants representing the pterophytes and gymnosperms. The PNP-B coding region could also not be amplified from the RNA isolated from the other plants representing angiosperms in this study.

After the confirmation of the PNP-B gene in the in the three plant species [wheat (betta), wheat (tugella) and maize (early pearl)], the amplified product was cloned to enable sequencing. For cloning the pDrive cloning vector was used. Therefore, the amplified PNP-B genes from the different plants were ligated into the pDrive cloning vector. Due to the presence of U/A overhangs, the PNP-B gene was directly inserted into the vector. The pDrive-PNP-B construct was then transformed into competent TOP10F` cells. Ten colonies resulting from the transformation procedure, for all the plants were screened for the pDrive-PNP-B construct by colony PCR. The colony PCR products were analysed by 1.0% agarose electrophoresis, as presented in figure 5.4 below. Only the colony PCR results for wheat (betta) are shown in the figure below.

(a)



(b)

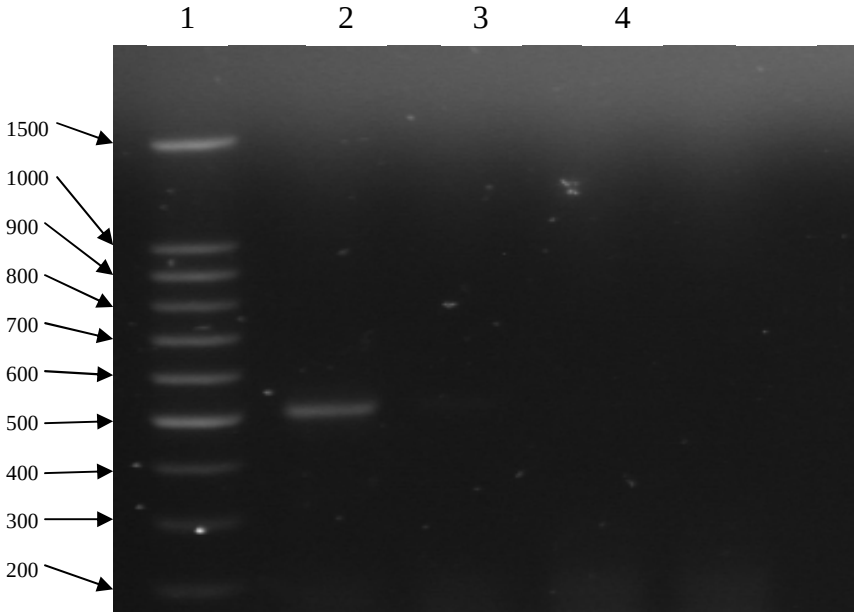


Figure 5.3 0.8% agarose gel electrophoresis of the RT-PCR amplification of PNP-B.

(a) Lanes: 1 –100 bp DNA Ladder (Promega); 2 – Wheat (Betta); 3 – Wheat (Tugella); 4 – Moss; 5 – Negative control (No RNA template).

(b) Lanes: 1 – 100 bp DNA Ladder (Promega); 2 – Maize (Early pearl); 3 – Cycad; 4 – Control (No RNA template)

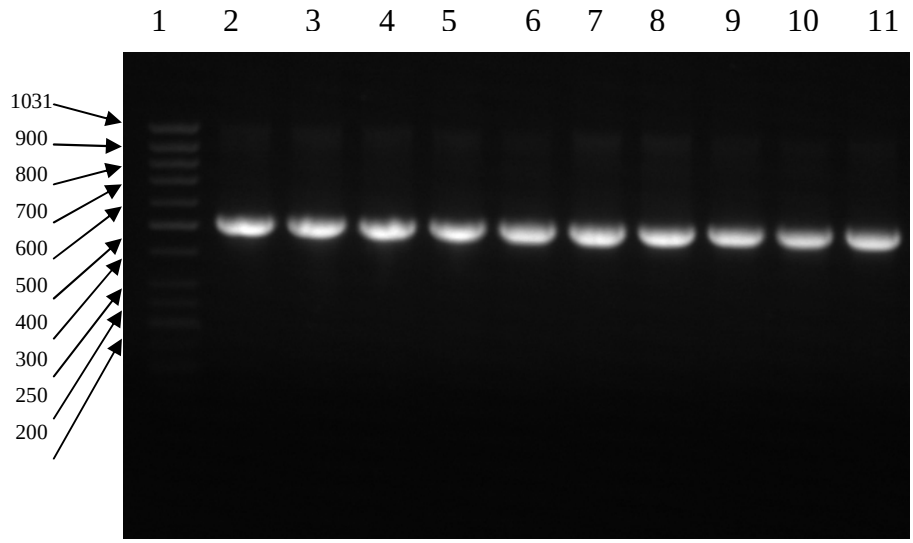


Figure 5.4 Verification of the pDrive-WhBPNP-B ligation by colony PCR, as visualised by 1.0% agarose electrophoresis. Lanes: 1 – O'GeneRuler™ 50 bp DNA Ladder; Colony PCR products from; 2 – colony 1; 3 – colony 2; 4 – colony 3; 5 – colony 4; 6- colony 5; 7 – colony 6; 8 – colony 7; 9 – colony 8; 10 – colony 9; 11 – colony 10.

A band corresponding to 510 bp was expected following colony PCR, (figure 5.4). For all the plants, colonies were identified that showed the presence of the pDrive-PNP-B construct. After the confirmation of the insert in the cloning vector, only one of the ten colonies was picked for the mini-preparation of the PNP-B plasmid for sequencing. The sequences of the plants where the PNP-B gene could be amplified are shown in figure 5.5 below (original chromatograms are attached in Appendix 1).

(a)

MSKSIVFFSTVLVFLFSFSYATPGIATFYTSYTRVTEVLKKERRSLQRVIHYGTMVEFVAKCSPRNAADLVTPCLTLARGNPRRSRS
LTIVSVAVLLRSIFLGKLLLRSLILSLGLLTLIIFRK

(b)

NKYKQQTYLIHFHINLQRNVIKSFSFISCLIASACYRGTQEGVMIAAASDTLWDNGRVCGKMFTVKCSGPRNAVPHPCTGKSAKVKI
VDHCPSGCASTLDLSREAFAQIANPVAGIINIDYFP

(c)

MSKSIVFFSTDYEHRLFSVRALQSEFVDKLLEPRRLAAHPPFASWRNSEEARTDRPSQQLRSLNGEWKLQALIFCENSRYIFVKSAH
FLTNRPKSAKSLINQKNRPRSGRVLFQFGTRVHYQRT

Figure 5.5 PNP-B sequences of the cereal plants. (a) Maize (Early pearl) (b) Wheat (Betta) (c) Wheat (Tugella).

Homology searches to the sequences obtained, using BLAST (Basic Local Alignment Search Tool; Altschul *et al.,* 1990) identified mostly the expansin family with significant sequence similarity to AtPNP-B. CjBAp12, citrus blight protein was also shown to have some sequence similarity. It was interesting to note that the only sequences that could be amplified using the AtPNP-B primer sequences were limited to cereals. It was therefore decided to attempt to identify possible other plant PNP-B sequences deposited in protein databases, using a signature pattern derived from these sequences. Pattern searches against PROSITE (http://au.expasy.org/prosite/) showed that AtPNP-B shares a structural motif with an endoglucanase belonging to family 45. The signature pattern ([PA]-C-[YN]-[RG]-[GFY]-[TEK]-[QDN]-[EQD]-G-[VT]-M-I-A-A-A-S-[DY]-[TVA]-[LFI]-W-[DN]-[NG]-G-[RA]-[VA]-C-[GN]) was derived from within the ANP-homologous region of AtPNP-B (33 - 66). Using this signature pattern however, only identified other proteins belonging to the endoglucanase 45 family.

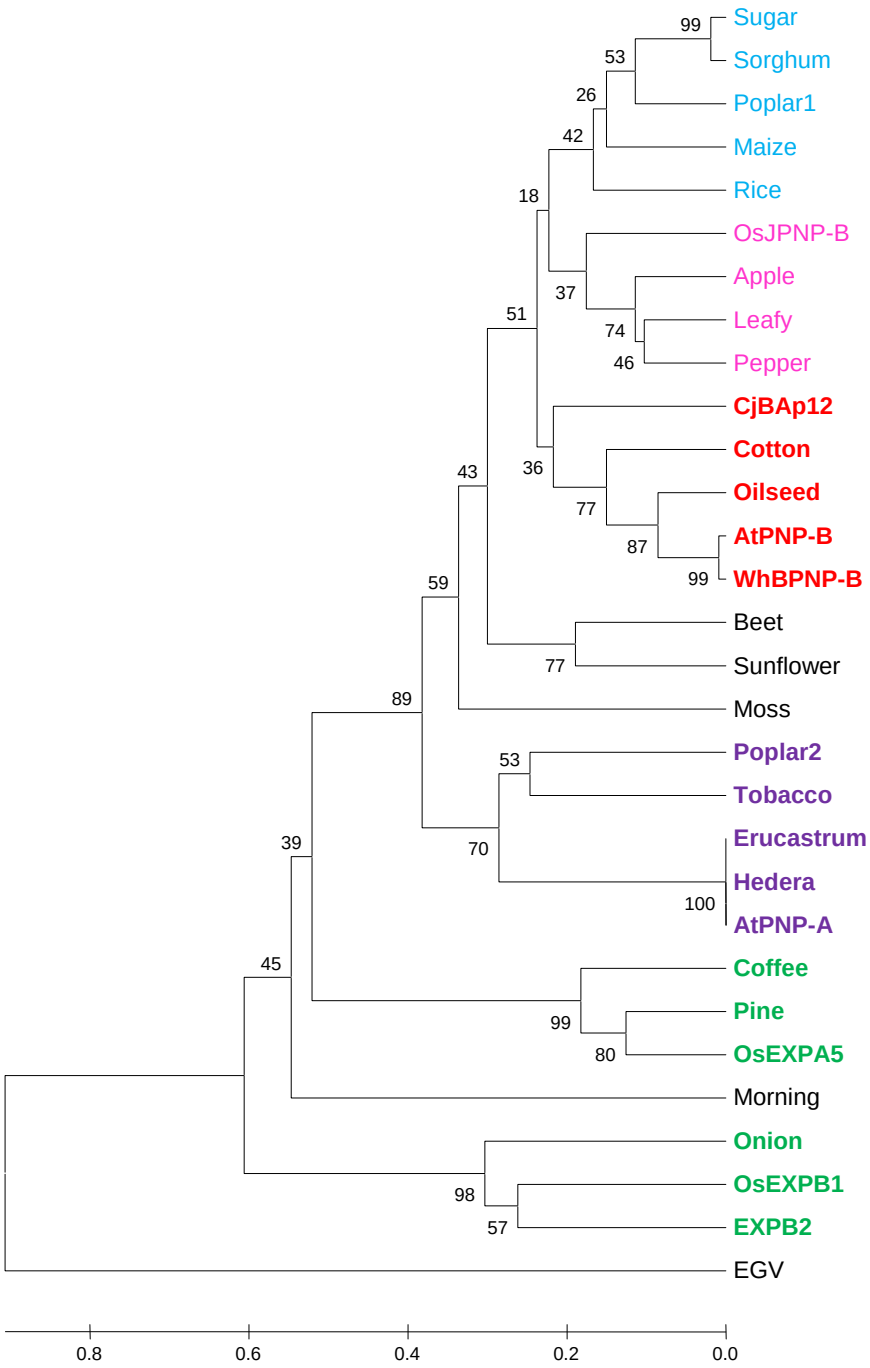


Figure 5.6 Phylogenetic tree of PNP-B like sequences encoded by various plant genomes. The numbers represent bootstrap values from 1 000 replicates. Sequence identification: Apple (TC10919), AtPNP-A (Q9ZV52.2), AtPNP-B (Q9M0C2.1), Beet (TC6463), CjBAp12 (Q9ZP41.1), Coffee (EE192624), Cotton (EX166047), Erucastrum (AAO85279), Hedera (AAM18791), Leafy Spurge (TC6438), Maize (TC418589), Morning glory (TC5961), Moss (NP13134739), Oilseed Rape (TC81032), Onion (CF451209), OsJPNP-B (BAD2245.1), Pepper (TC11504), Pine (DR019211), Poplar (TC114006), Rice (TC360733), Sugarcane (CA286353), Sorghum (CF484893), Sunflower (TC33321), Tobacco (EB424615), WhBPNP-B, OsEXPA5 (AAF62180), OsEXPB1 (AAF72983), EXPB2 (AAB61710), EGV (P43316). The AtPNP-B cluster is shown in red and the AtPNP-B related cluster 1 is shown in pink while cluster 2 is shown in blue. The AtPNP-A cluster is shown in purple and the expansins cluster in green.

An alternative approach was therefore followed to identify other possible gene sequences encoding for PNP-B like proteins from different plant genomes using BLAST (tblastn and tblastx). The search was done against all the plant genomes that are available on the "Gene Index Project" (http://compbio.dfci.harvard.edu/tgi/plant.html) and seventeen genomes gave sequences that were related to PNP-B. The conserved MIAAA sequence was used to distinguish between protein sequences that are related to PNP-B and those related to PNP-A. The PNP-B sequences that were obtained were then aligned and a phylogenetic tree was constructed as shown in figure 5.6 above.

5.4 Discussion

The intron predicted in the *Arabidopsis* gene for PNP-B was found in the cloned sequences [wheat (betta and tugella) and maize (early pearl)]. This then raises the question of whether this is truly an intron or not, or maybe an insertion as the PNP-B coding region was amplified from RNA. A DNA contamination test was done on the isolated RNA to eliminate the possibility of contaminating DNA in the RNA sample. The DNA contamination test proved negative, showing that there was no DNA contamination (data not shown). The presence of the 138 bp fragment within the PNP-B gene was contrary to expectations since if it was an intron, it would have been spliced out in the mature RNA. Thus, this suggests that PNP-B could be a pseudogene based on the plants that had introns in them.

Pseudogenes are DNA sequences that resemble functional genes but are thought to have no purpose. The first pseudogene was reported in 1977 (Jacq *et al.,* 1977) and pseudogenes have since been grouped into two namely; 'processed' and 'unprocessed' pseudogenes (Gibson, 1994). 'Processed' pseudogenes lack introns and certain regulator genes, often terminate in adenine series and are flanked by direct repeats, which are associated with movable genetic elements. They may be complete or incomplete copies of genes or mixtures of several genes (Gibson, 1994). While 'unprocessed' pseudogenes have introns and associated regulatory sequences. Their expression is usually prevented by a misplaced stop codon or codons. They are believed to have arisen by gene duplication, which produced an extra copy of the functional gene. The extra copy could then accumulate mutations without harming the organism since it would still have a completely functional original copy (Gibson, 1994).

Therefore, if PNP-B is a pseudogene at all, this makes it an 'unprocessed' pseuodogene since it has an intron and the associated regulatory sequences. Stop codons were also identified within the PNP-B gene after sequencing thus, supporting the fact that 'unprocessed' pseudogenes usually have misplaced stop codons. The misplaced stop codons prevent expression and therefore explain why the expression of the PNP-B gene that was attempted in Chapter 4 was not successful. It can also be speculated that if PNP-B is a pseudogene then its functional gene is PNP-A.

However, although based on the cloned sequences, we are suggesting that PNP-B could be a pseudogene; the evolutionary gene duplication hypothesis also suggests that over time, random mutations may produce a new gene with new functions by using the gene duplicate while maintaining the original gene function (Long & Langley, 1993). Thus, it is also possible that if PNP-B is not functional at the moment then over time it will obtain a new function that is linked to the original gene function. Therefore, it can be speculated that since PNP-B has got sequence

similarity to PNP-A and CjBAp12 then it could have a role in pathogenesis response to stresses that affect water and salt homeostasis.

The PNP-B gene was only amplified and sequenced from wheat (betta and tugella) and maize (early pearl) while it could not be amplified from the other plants where RNA was isolated. The reason why the gene could not be amplified from the other plants was due to the diversity in the 5` region of the PNP-B of the genes in different plants. This would result in the forward primers that were used in this study, which were designed based on the *Arabidopsis thaliana* gene sequence, not annealing to the template RNA. This is evidenced by the homology searches that were done using BLAST against different plant genomes using the AtPNP-B sequence as a query. Examples of plants from which PNP-B related molecules were identified and from which PNP-B amplification was also attempted but was unsuccessful were tobacco, sorghum and moss. For moss the only region of the AtPNP-B gene that was identified was from 73 - 230, for sorghum it was 133 – 362 and for tobacco it was the region from 73 – 338. Therefore, since the primers were designed against the first about 25 bases of AtPNP-B, the gene could not be amplified from the other plants.

Sequences from the BLAST searches that were done against the different plant genomes were aligned and a phylogenetic tree was constructed. The phylogenetic tree showed that AtPNP-B clustered with CjBAp12 in a group that contained other PNP-B related molecules from other plants such as cotton, oilseed and wheat (betta). There were also two other clusters in the phylogenetic tree (shown in blue and pink in figure 5.6 above) that were shown to be related to PNP-B. The cluster shown in blue represented cereals (monocotyledons) namely; rice, maize, sorghum and sugarcane. The cluster shown in pink represented dicotyledons (apple, leafy spurge and pepper), with the only anomaly being the rice sequence. Another cluster group contained

AtPNP-A. This showed that AtPNP-B is more evolutionarily related to CjBAp12 than it is to AtPNP-A.

From the blast (tblastx) searches against the different plant geneomes showed us that there are some plants in which the intron is not included such as; rice (oryza sativa), leafy spurge, oilseed rape, onion, poplar, sugar cane, sunflower and tobacco. There were also other plants where only the region from 200 – 500 was the only one that could be identified from the AtPNP-B gene that was used as the query. These plants included apple, barley, coffee, cotton, maize (zea mays), morning glory, moss, pine and sorghum. There is a possibility that these plants could be lacking the intron as well.

Homology searches using BLAST in NCBI identified mostly expansins as the molecules that are related to PNP-B. This was consistent with the results of Ludidi *et al.*, (2002) where they also identified expansins as molecules with similarity to irPNP-like molecules. This is also consistent with the results from the phylogenetic tree which showed that the PNP-B related proteins from onion, pine and coffee were clustering in the same group with the α- and β- expansins. PNPs are similar to expansins but do not contain the tryptophan and tyrosine rich C-terminal putative polysaccharide-binding domain typical of expansins, thus, making PNPs considerably shorter than the expansin sequences (~ 12 – 14 kDa rather than ~ 25 kDa). The C-terminal in expansins that is absent in PNPs was assigned as a putative cellulose binding, and thus cell wall binding region of expansins (Cosgrove, 1999). The absence of the expansin C-terminal in PNP-Bs shows that although, PNP-Bs could be related to expansins; they do not interact with the cell wall and are therefore mobile. It has been previously demonstrated that irPNPs act on protoplasts (plant cells without cell walls) as well as microsomes (Gehring *et al.*, 1996; Pharmawati *et al.*, 2001), indicating that the novel peptides specifically interact with the plasma membrane. This shows

that the cell wall cannot be an obligatory substrate for PNP-Bs. Thus, the absence of the cellulose binding C-terminal of expansins, in PNP-Bs points the molecules to having a systemic role in water and solute homeostasis. Furthermore to show that PNP-B does not act on the cell wall like expansins; the BLAST homology searches also identified CjBAp12, a blight-induced protein which has no apparent expansin activity (Ceccardi *et al.,* 1998). Thus, it can be speculated that PNP-B has a role in pathogenesis response which may involve regulation of water and salt homeostasis.

Expansins also share structural motifs with the putative catalytic sites of endoglucanases of the family 45, which in turn do not have expansin activity (Cosgrove, 1999; Cosgrove, 2000). Thus, pattern searches against PROSITE showed that PNP-Bs share a structural motif with an endoglucanase belonging to family 45. The fact that expansins, endoglucanase family 45 and PNP-Bs share structural motifs reflects an evolutionary relationship among the molecules and therefore a relationship in function where the structure of PNP-B had to be changed to adapt to a new function that is different from both expansins and endoglucanases. Unfortunately further studies will have to be conducted to elucidate the exact function of PNP-B in plants.

5.5 Conclusion

In conclusion it can be noted that there is a possibility that PNP-B could be a pseudogene only in certain plants that contain the intron in them such as [*Arabidopsis*, wheat (betta and tugella) and maize (early pearl)] and its functional gene is PNP-A, which has been extensively studied on and has evidence pointing towards a role in water and salt homeostasis. But in other plants such as rice, leafy spurge, oilseed rape, onion, poplar, sugar cane, sunflower and tobacco it appears to be

a functional protein since there is no intron in those particular plants. In this case, PNP-B would take up a function related to water and salt homeostasis.

It can also be noted that PNP-B is evolutionarily related to expansins and the endoglucanase family 45, and that PNP-Bs lost the C-terminal domain important for the functional role of expansins and endoglucanases. The loss of the domain will then relate PNP-B to a systemic role and also increased mobility. Increased mobility is supported by the fact that CjBAp12, a citrus blight-induced protein which has sequence similarity (54%) to AtPNP-B is also a systemically mobile protein (Ceccardi *et al.,* 1998), present, but not synthesized in leaves.

Furthermore, given the evolutionary relationship of PNP-Bs and expansins we can go with the argument of Ludidi *et al.,* (2002) which gives a possible link in function of expansins and PNPs. There has been evidence that PNP molecules can cause enhanced osmoticum-dependent water uptake (Maryani *et al.,* 2001; Pharmawati *et al.,* 2001) and hence increases in cell turgor. Increasing cell turgor will pose stress on cell walls and may in turn signal the need for expansin-dependent wall loosening. Thus, inferring from this it is possible that PNP-B and expansins both respond towards water and salt homeostasis but in different capacities. Where PNP-B would enhance water uptake and expansins will be involved in cell wall loosening.

This work has revealed some evidence pointing PNP-Bs towards a functional role in both abiotic stress and pathogen response. However, further work needs to be done to conclusively determine whether PNP-Bs are functional proteins that have a role in water and salt homeostasis or a possible role in pathogenesis response.

CHAPTER SIX: GENERAL DISCUSSION & FUTURE RESEARCH WORK

6.1 General Discussion & Conclusion

Sustaining water and solute homeostasis is a key requirement for living systems and in vertebrates homeostasis is, in part achieved by natriuretic peptides (NPs), a family of peptide hormones (Anand-Srivastava & Trachte, 1993; Kone, 2001; Suzuki *et al.,* 2001). There is structural and functional evidence to suggest that an immunological related peptide hormone system may operate in plants (Vesely *et al.,* 1993; Billington *et al.,* 1997; Gehring, 1999). Biologically active plant natriuretic peptide immune-analogues (irPNP) have been isolated and purified by immunoaffinity chromatography (Billington *et al.,* 1997). Two *Arabidopsis thaliana* irPNP-encoding genes termed AtPNP-A and AtPNP-B have also been identified and isolated (Ludidi *et al.,* 2002).

This study was done with a major focus on PNP-B since PNP-A has been extensively studied. Immunological evidence for the presence of PNP-like molecules has been found in angiosperm species (Vesely *et al.,* 1993; Billington *et al.,* 1997; Gehring, 1999; Ludidi *et al.,* 2002). In this study the coding region of the PNP-B gene was successfully amplified from wheat (tugella and betta) and maize (early pearl) (Chapter 5) and RT-PCR was used to amplify the gene. A 138 bp intron was identified for all the plants where the PNP-B gene was amplified, including from *Arabidopsis thaliana* in chapter 4. The presence of the intron after amplification of the PNP-B coding region from template RNA from *Arabidopsis*, wheat (betta and tugella) and maize (early pearl) suggested that PNP-B could be an ‘unprocessed’ pseudogene. ‘Unprocessed’ pseudogenes have misplaced stop codons which result in them not being expressed. Thus, the attempts at

amplifying PNP-B in chapter 4 were not successful. It was also speculated that PNP-A could be the functional gene of PNP-B.

Although it would appear that PNP-B has an intron we found from searching the plant genomes by tblastx that some did not have an intron region as in, rice, leafy spurge, oilseed rape, onion, poplar, sugar cane, sunflower and tobacco and would therefore encode a functional PNP-B protein. Bioinformatic and structural analysis (chapter 3) as well as evolution studies (chapter 5) pointed towards PNP-Bs having a functional role in both abiotic stress and pathogen responses. Firstly, AtPNP-B was found to have some sequence similarity with AtPNP-A (chapter 3), which has been extensively studied. Biologically active PNP immuno-analogues (irPNP) and recombinant AtPNP-A have been shown to have the following effects on plants;

- Induction of stomatal opening in a concentration- and conformation dependent manner (Gehring *et al.*, 1996; Pharmawati *et al.*, 1998a, 2001; Maryani *et al.*, 2001; Morse *et al.*, 2004),
- Enhancement of osmoticum-dependent volume changes in leaf mesophyll protoplasts (Maryani *et al.*, 2001),
- Modulation of ion fluxes across plant membranes (Pharmawati *et al.*, 1999) and
- Induction of volume changes in protoplasts isolated from Arabidopsis cell suspensions (Rafudeen *et al.*, 2003; Morse *et al.*, 2004).

Furthermore it has also been reported that in the brassicaceous weed *Erucastrum strigosum* the PNP-like molecules are significantly up-regulated under salinity stress conditions (Rafudeen *et al.*, 2003). The PNP-like molecules have also been significantly up-regulated in *Arabidopsis thaliana* suspension culture cells in response to 150 mM NaCl and in response to iso-molar

amounts of sorbitol (Rafudeen *et al.,* 2003). In addition, time-dependent increases in PNP in the xylem of the African sage *Plectranthus ciliatus* were registered after shoot removal (Maryani *et al.,* 2003) which causes severe homeostatic disturbances. Taken together, these observations establish PNP-A as having a role in osmotic stress-responsiveness and the regulation of water and solute movements. Thus, it can be inferred that since PNP-B has sequence similarity with PNP-A, it might also have a function in water and salt homeostasis.

Secondly, the transcription factor binding sites (TFBS) that were predicted upstream of the AtPNP-B and OsJPNP-B in chapter 3, also showed that PNP-Bs might have a role in water and salt homeostasis. The TFBS that were predicted for both AtPNP-B and OsJPNP-B were mostly those that were regulating water stress and homeostasis processes.

However, the sequence similarity of AtPNP-B to CjBAp12 (Chapter 3) and the fact that AtPNP-B and CjBAp12 clustered in the same group in the phylogenetic tree (Chapter 5) suggested a pathogenesis-related role for PNP-B. This was further supported by the prediction of TFBS (Chapter 3) that were related to biotic stress upstream of both the AtPNP-B and OsJPNP-B genes. This evidence suggests that there could be an overlap of roles for PNP-B where PNP-B responds to biotic stresses that are related to water and salt homeostasis in plants.

To further characterize PNP-B we went on in chapter 5 to search for other known proteins that are related to it. Using signature patterns searches, the protein was shown to have an evolutionary link with endoglucanase family 45 molecules. When BLAST was used to search the protein databases for more molecules that could be related to PNP-B, mostly expansins were identified. This was consistent with previous studies that have shown that expansins share structural motifs

with the putative catalytic sites of endoglucanases of the family 45 (Cosgrove, 1999). Thus, this shows us that there is an evolutionary link among the three molecules.

PNPs might be similar to expansins but they do not contain the C-terminal putative polysaccharide-binding domain typical of expansins and glucanases, thus, making PNPs considerably shorter. This loss in the C-terminal suggests that during evolution the domain was lost so that PNP-B could adopt another function other than the function of expansins and glucanases in plants. This suggests a systemic role for PNP-Bs. The systemic role is further supported by the sequence similarity that AtPNP-B has with CjBAp12 and the fact that AtPNP-B and CjBAp12 cluster together in the same group in the phylogenetic tree. CjBAp12 appears to be a systemically mobile protein (Ceccardi *et al.,* 1998) since it is present but not synthesized in the leaves.

The relationship that PNP-B has with expansins also suggests a divergence in the substrate that the molecules act on. Expansins are proteins that promote cell wall loosening and extension (Cosgrove, 1999) thus they act on the cell wall while PNPs have been shown to act on the membrane. However, irPNP-dependent processes have been observed in experimental systems that do not contain the cell wall such as protoplasts or microsomal and plasma membrane vesicles (Suwastika *et al.,* 2000; Maryani *et al.,* 2000; Pharmawati *et al.,* 2001; Maryani *et al.,* 2001). This suggests that PNP-B could also act directly on the plasma membrane.

Another BLAST search against the different plant genomes in the “Gene Index Project” database in chapter 5 also showed that PNP-B related molecules were also present in other plants. The identification of other PNP-B related molecules could help explain why PNP-B could not be amplified from various plants other than the three cereal plants. This was due to the diversity of

the signal peptide coding region of the different PNP-B related genes, the primers that were designed based on the PNP-B sequence from *Arabidopsis thaliana* could not amplify the gene from some of the plants. The sequences that were obtained from other plant genomes were also used in the construction of a phylogenetic tree.

In conclusion, it was shown that the PNP-B gene is present in other plants from where the gene could not be amplified thus there is a need to design degenerate primers so as to amplify the respective PNP-B genes from the different plants. Due to the sequence similarity of AtPNP-B to both AtPNP-A and CjBAp12, it can be concluded that PNP-B may have a function in pathogenesis related stresses that affect water and salt homeostasis. The evolutionary relationship of PNP-B with expansins and endoglucanases also suggests a divergence in function where PNPs are systemic mobile and act on cell membranes, unlike expansins which act on the cell walls.

6.2 Future Research Work

In this study, RT-PCR was done on total RNA from 15 different plants but the PNP-B gene was successfully amplified from only three plants. Thus for future studies it is important to design degenerate primers or specific primers, that will be able to amplify the gene from all the plants where the gene can be found. After amplification of the gene from different plants, it is also important to amplify and express a PNP-B gene lacking an intron. Expression of PNP-B can then be optimised using these plants. The purified recombinant PNP-B will be used to investigate the cellular responses of plants to recombinant PNP-B and thus elucidate its function. The recombinant PNP-B will also be used to further analyse the structure of PNP-B using experimental methods such as NMR (nuclear magnetic resonance) and X-ray crystallography.

It has been reported that the region sufficient for triggering protoplast swelling and stomatal guard cell opening of AtPNP-A lies within the region 33 – 66 which has been shown to be homologous to ANP (Morse *et al.,* 2004; Wang *et al.,* 2007). Thus, it is also important to confirm the region of PNP-B that is responsible for its function.

A further aspect of this work would be to investigate the responses of PNP-B to biotic and abiotic plant stress responses. Immuno-localisation of PNP-B in the plants pre- and post stress is also suggested in order to have an idea of the localization of PNP-B in plants. Furthermore, it will be important to research on the regulatory elements of PNP-B using DNA foot-printing and reporter genes such as GUS.

Further research work on PNP-B will help to bring about a conclusive role for the molecule. The functional characterization of PNP-B will help to improve biotechnological stress tolerance strategies and to further elucidate the role of proteins/peptides involved in plant stress signaling responses.

REFERENCES

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2003). "*Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling." *Plant Cell,* 15: 63-78.

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwaski, T., Hosokawa, D. and Shinozaki, K. (1997). "Role of *Arabidopsis* MYC and MYB homologs in drought- and abscisic acid-regulated gene expression." *Plant Cell,* 9: 1859-1868.

Albrigo, L. G. & Young, R. H. (1981). "Phloem zinc accumulation in citrus trees affected with blight." *Hort Science,* 16: 158-160.

Aldington, S. & Fry, S.C. (1993). "Oligosaccharins." *Advances in Botanical Research,* 19: 1–107.

Alexandrov, D. A. (1994). Filipchenko and Dobzhansky: Issues in Evolutionary Genetics in the 1920s. In *The Evolution of Theodosius Dobzhansky*, ed. MB Adams, Princeton: Princeton University Press.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). "Basic local alignment search tool." *Journal of Molecular Biology,* 215:403-10.

Alvarez, M. E. (2000). "Salicylic acid in the machinery of hypersensitive cell death and disease resistance." *Plant Molecular Biology,* 44: 429-442.

Amin, J., Carretero, I. A. and Ito, S. (1996). "Mechanisms of action of atrial natriuretic factor and C-type natriuretic peptide." *Hypertension,* 27: 684–687.

Anand-Srivastava, M. B. & Trachte, G. J. (1993). "Atrial natriuretic factor receptor and signal transduction mechanisms." *Pharmacological Reviews,* 45:455–497.

Antikainen, M., Griffith, M., Zhang, J., Hon, W-C., Yang, D. S. C. and Pihakashi-Maunsbach, K. (1996). "Immunolocalization of antifreeze proteins in winter rye leaves, crowns, and roots by tissue printing." *Plant Physiology*, 110: 845-857

Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006). "The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling." *Bioinformatics*, 22: 195-201.

Arteca, R. N. (1996). "Plant Growth Substances: Principles and Applications." Chapman & Hall, New York, pp332.

Atıcı, O. & Nalbantoglu, B. (2003). "Antifreeze proteins in higher plants." *Phytochemistry*, 64: 1187–1196.

Baker, S. S., Wilhelm, K. S. and Thomashow, M. F. (1994). "The 5`-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression." *Plant Molecular Biology*, 24: 710-713.

Bandurski, R. S., Cohen, J. D., Slovin, J. and Reinecke, D. M. (1995). "Auxin biosynthesis and metabolism." In Davies, P. J. (ed). Plant Hormones – Physiology, Biochemistry and Molecular Biology, Kluwer Academic Publishers: Dordrecht, Boston, London. pp. 39-65.

Baneyx, F. (1999). "Recombinant protein expression in *Escherichia coli*." *Current Opinion in Biotechnology*, 10: 411–421.

Bartel, B. (1997). "Auxin biosynthesis." *Annual Review of Plant Physiology & Plant Molecular Biology*, 48: 51–66.

Berger, S., Bell, E. and Mullet, J. E. (1996). "Two methyl jasmonste-insensitive mutants show altered expression of Atvsp in response to methyl jasmonate and wounding." *Plant Physiology*, 111: 525-531.

Bergeron, B. (2003). "The complete, practical guide to bioinformatics for life scientists: Bioinformatics Computing." Pearson Education, Inc.

Bergey, D. R., Howe, G. A. and Ryan, C. A. (1996). "Polypeptide signaling for plant defensive genes exhibits analogies to defense signaling in animals" *Proceedings of the National Academy of Sciences USA,* 93: 12053–12058.

Bernier, F. & Berna, A. (2001). "Germins and germin-like proteins: plant do-all proteins. But what do they do exactly?" Plant Physiology & Biochemistry, 39:545–554.

Bewley, J. D. (1997). "Seed germination and dormancy" *Plant Cell,* 9: 1055–1066.

Bialek, K. & Cohen, J. D. (1986)."Isolation and partial characterization of the major amide-linked conjugate of indole-3-acetic acid from *Phaseolus vulgaris L*." *Plant Physiology,* 80: 99-104.

Billington, T., Pharmawati, M. and Gehring, C. A. (1997). "Isolation and immunoaffinity purification of biologically active plant natriuretic peptide." *Biochememical & Biophysical Research Communications,* 235: 722-725.

Bishop, G. J. & Yokota, T. (2001). "Plant steroid hormones, brassinosteroids: Current highlights of molecular aspects on their synthesis/metabolism, transport, perception and response." *Plant Cell Physiology,* 42(2): 114-120.

Bohlmann, H. (1994). "The role of thionins in plant protection." *Critical Reviews in Plant Sciences,* 13: 1–16.

Bol, J. F., Linthorst, H. J. M. and Cornelissen, B. J. C. (1990). "Plant pathogenesis-related proteins induced by virus infection." *Annual Review of Phytopathology,* 28: 113-138.

Bolter, C. & Jongsma, M. A. (1997). "The adaptation of insects to plant protease inhibitors." *Journal of Insect Physiology,* 43: 885-895.

Borsing, L., Berger, E. G. and Malissard, M. (1997). "Expression and purification of His-tagged beta-1,4 galactosyltransferase in yeast and in COS cells." *Biochemical & Biophysical Research Communications,* 240: 586–590.

Boudart, G., Jamet, E., Rossignol, M., Lafitte, C., Borderies, G., Jauneau, A., Esquerre-Tugaye, M. T. and Pont-Lezica, R. (2004). "Cell wall proteins in apoplastic fluids of Arabidopsis thaliana rosettes: Identification by mass spectrometry and bioinformatics." *Proteomics,* 5: 212-221.

Bourne, P. E. & Weissig, H. (2003). "Structural Bioinformatics." John Wiley & Sons, Inc.

Broadway, R. M. & Duffey, S. S. (1986). "Plant proteinase inhibitors mechanism of action and effect on the growth and digestive physiology of larval *Heliothis zea* and *Spodoptera exiqua*." *Journal of Insect Physiology,* 32: 827-833.

Broekaert, W. F., Terras, F. R. G. and Cammue, B. P. A. (2000). "Induced and preformed antimicrobial proteins." *Mechanisms of Resistance to Plant Diseases,* Dordrecht: Kluwer. pp. 371–477.

Brown, R. L., Kazan, K., McGrath, K. C., Maclean, D. J. and Manners, J. M. (2003). "A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of *Arabidopsis*." *Plant Physiology,* 132: 1020-1032.

Brummell, D. A., Harpster, M. H., Civello, P. M., Palys, J. M., Bennett, A. B. and Dunsmuir, P. (1999). "Modification of expansin protein abundance in tomato fruit alters softening and cell wall polymer metabolism during ripening." *Plant Cell,* 11: 2203–2216.

Brunner, F., Stintzi, A., Fritig, B. and Legrand, M. (1998). "Substrate specificities of tobacco chitinases." *The Plant Journal,* 14: 225–234.

Buchanan-Wollaston, V., Earl, S., Harrison, E., Mathas, E. and Navabpour, S. (2003). "The molecular analysis of leaf senescence—a genomics approach." *Plant Biotechnology Journal,* 1: 3–22.

Bufe, A., Spangfort, M. D., Kahlert, H., Schlaak, M. and Becker, W-M. (1996). "The major birch pollen allergen, Bet v 1, shows ribonuclease activity." *Planta*, 199: 413–415.

Buhtz, A., Kolasa, A., Arlt, K., Walz, C. and Kehr, J. (2004). "Xylem sap protein composition is conserved among different plant species." *Planta*, 219: 610–618.

Bujard, H., Gentz, R., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M. T. and Dobberstein, B. (1987). "A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo." *Methods in Enzymology*, 155: 416-433.

Busk, P. K. & Pages, M. (1998). "Regulation of abscisic acid-induced transcription." *Plant Molecular Biology*, 37: 425-435.

Butenko, M. A., Stenvik, G.-E., Alm, V., Saether, B., Patterson, S, E. and Aalen, R. B. (2006). "Ethylene-dependent and –independent pathways controlling floral abscission are revealed to converge using promoter::reporter gene constructs in the ida abscission mutant." *Journal of Experimental Botany*, 57: 3627-3637.

Butenko, M. A., Patterson, S, E., Grini, P. E., Stenvik, G. E., Amundsen, S. S., Mandal, A. and Aalen, R. B. (2003). "Inflorescence deficient in abscission controls floral organ abscission in *Arabidopsis* and identifies a novel family of putative ligands in plants." *Plant Cell*, 15: 2296-2307.

Calos, M. P. (1978). "DNA sequence for a low-level promoter of the *lac* repressor gene and an 'up' promoter mutation." *Nature*, 274: 762–765.

Cameron, R. K., Dixon, R. A. and Lamb, C. J. (1994). "Biologically induced systemic acquired resistance in *Arabidopsis thaliana*." *Plant Journal*, 5: 715-725.

Campalans, A., Kondorosi, A. and Crespi, M. (2004). "Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*." *Plant Cell*, 16: 1047-1059.

Carmelo, C., Lopez-Farre, A., Riesco, A., Olivera, A., Cragoe, E. J., Tsai, P., Briner, V. A. and Schrier, R. W. (1994). "Atrial Natriuretic and cGMP inhibit Na^+/H^+ antiporter in vascular smooth muscle cell culture." *Kidney International,* 45: 66-75.

Carpita, N. C. (1996). "Structure and biogenesis of the cell walls of grasses." *Annual Review of Plant Physiology & Plant Molecular Biology,* 47: 445–476.

Casson, S. A., Chilley, P. M., Topping, J. F., Evans, I. M., Souter, M. A. and Lindsey, K. (2002). "The POLARIS gene of *Arabidopsis* encodes a predicted peptide required for correct root growth and leaf vascular patterning." *Plant Cell,* 14: 1705-1721.

Castillo, R. M., Mizuguchi, K., Dharanaj, V., Albert, A., Blundell, T. L. and Murzin, A. G. (1999). "A six-stranded double-psi β barrel is shared by several protein superfamilies." *Structure,* 7: 227-236.

Ceccardi, T. L., Barthe, G. A. and Derrick, K. S. (1998). "A novel protein associated with citrus blight has sequence similarities to expansin." *Plant Molecular Biology,* 38: 775-783.

Cercos, M., Gomez-Cadenas, A. and Ho, T. H. D. (1999). "Hormonal regulation of a cysteine proteinase gene, EPB-1, in barley aleurone layers: cis- and trans-acting elements involved in the co-ordinated gene expression regulated by gibberellins and abscisic acid." *Plant Journal,* 19: 107-118.

Cermak, R., Kleta, R., Forssmann, W. G. and Schlatter, E. (1996). "Natriuretic peptides increase a K^+ conductance in rat mesangial cells." *European Journal of Physiology,* 431: 571-577.

Charon, C., Sousa, C., Crespi, M. and Kondorosi, A. (1999). "Alteration of enod40 expression modifies *Medicago truncatula* root nodule development induced by *Sinorhizobium meliloti*." *Plant cell,* 11: 1953-1966.

Charon, C., Johansson, C., Kondorosi, E., Kondorosi, A. and Crespi, M. (1997). "Enod 40 induces dedifferentiation and division of root cortical cells in legumes." *Proceedings of the National Academy of Sciences USA*, 94: 8901-8906.

Chart, H., Smith, H. R., La Ragione, R. M. and Woodward M. J. (2000). "An investigation into the pathogenic properties of *Escherichia coli* strains BLR, BL21, DH5alpha and EQ1." *Journal of Applied Microbiology*, 89: 1048–1058.

Chen, B. P. & Hai, T. (1994). "Expression vectors for affinity purification and radiolabeling of proteins using *Escherichia coli* as host." *Gene*, 139: 73–75.

Chilley, P. M., Casson, S. A., Tarkowski, P., Hawkins, N., Wang, K. L., Hussey. P. J., Beale, M., Ecker, J. R., Sandberg, G. K. and Lindsey, K. (2006). "The POLARIS peptide of Arabidopsis regulates auxin transport and root growth via effects on ethylene signalling." *Plant Cell*, 18: 3058-3072.

Cho, H.-T. & Cosgrove, D. J. (2000). "Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*." *Proceedings of the National Academy of Sciences USA*, 97: 9783–9788.

Choi, D., Lee Y., Cho, H.-T. and Kende, H. (2003). "Regulation of expansin gene expression affects growth and development in transgenic rice plants." *Plant Cell*, 15: 1386–1398.

Chothia, C. & Lesk, A. (1986). "The relation between the divergence of sequence and structure in proteins." *EMBO Journal*, 5(4): 823-826.

Christensen, A. B., Ho Cho, B., Naesby, M., Gregersen, P. L. and Brandt, J. (2002). "The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins." *Molecular Plant Pathology*, 3:135–144.

Clark, S. E., Williams, R. W., and Meyerowitz, E. M. (1997). "The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*." *Cell*, 89: 575–585.

Cohen, J. D. (1982). "Identification and quantitative analysis of indole-3-acetyl-L-apartate from seeds of *Glycine max* L." *Plant Physiology*, 74: 749-753.

Cohen, M. (1974). "Diagnosis of young tree decline, blight and sand hill decline by measurement of water uptake using gravity injection." *Plant Disease Reporter,* 58: 801-805.

Cohen, J. D. & Bandurski, R. S. (1982). "Chemistry and physiology of the bound auxins." *Annual Review of Plant Physiology*, 33: 403-430.

Compaan, B., Ruttink, T., Albrecht, C., Meeley, R., Bisseling, T. and Franssen, H. (2003). "Identification and characterization of a *Zea mays* line carrying a transposon-tagged ENOD40." *Biochemical & Biophysical Acta.*, 629: 84-91.

Compaan, B., Yang, W-C., Bisseling, T. and Franssen, H. (2001). "ENOD40 expression in the pericycle precedes cortical cell division in *Rhizobium*-legume interaction and the highly conserved internal region of the gene does not encode a peptide." *Plant & Soil,* 230: 1-8.

Connell, S. R., Tracz, D. M., Nierhaus, K. H. and Taylor, D. E. (2003). "Ribosomal protection proteins and their mechanism of tetracycline resistance." *Antimicrobial Agents & Chemotherapy,* 47: 3675–3681.

Cooper, G. M. & Hausman, R. E. (2007). "The Cell: A molecular Approach." 4th Edition, ASM Press, Washington, D.C.

Cooper, J. B. & Long, S. R. (1994). "Morphogenetic rescue of *Rhizobium* meliloti nodulation mutants by trans-zeatin secretion." *Plant Cell,* 6: 215-225.

Cornelissen, B. J. C., Horowitz, J., Van Kan, J. A. L., Goldberg, R. B. and Bol, J. F. (1987). "Structure of tobacco genes encoding pathogenesis-related proteins from the PR-1 group." *Nucleic Acids Research,* 15:6799–6811.

Cosgrove, D. J. (2000). "Loosening of plant cell walls by expansins." *Nature,* 407: 321–326.

Cosgrove, D. J. (1999). "Enzymes and other agents that enhance cell wall extensibility." *Annual Review of Plant Physiology & Plant Molecular Biology,* 50: 391–417.

Cosgrove, D.J. (1997). "Relaxation in a high-stress environment: The molecular basis of extensible walls and cell enlargement." *Plant Cell,* 1031-1041.

Cosgrove, D. J., Li, L. C., Cho, H-T., Hoffmann-Benning, S., Moore, R. C. and Blecker, D. (2002). "The growing world of expansins." *Plant Cell Physiology*, 43(12): 1436-1444.

Creelman, R. A. & Mullet, J. E. (1997). "Oligosaccharins, brassinolides, and jasmonates: Nontraditional regulators of plant growth, development and gene expression." *Plant Cell,* 9: 1211-1223.

Custers, J. H. H. V., Harrison, S. J., Sela-Buurlage, M. B., Van Deventer, E. and Lageweg, W. (2004). "Isolation and characterization of a class of carbohydrate oxidases from higher plants, with a role in active defence." *Plant Journal*, 39: 147–160.

Davanloo, P., Rosenberg, A. H., Dunn, J. J. and Studier, F. W. (1984). "Cloning and expression of the gene for bacteriophage T7 RNA polymerase *Proceedings of the National Academy of Sciences USA,* 81: 2035-2039.

Davies, P. J. (1995). "The plant hormones: Their nature, occurrence and functions." In Davies, P. J. (ed). Plant Hormones – Physiology, Biochemistry and Molecular Biology. Kluwer Academic. Netherlands. pp. 1-12.

Davis, G. D., Elisee, C., Newham, D. M. and Harrison R. G. (1999). "New fusion protein systems designed to give soluble expression in *Escherichia coli*." *Biotechnology & Bioengineering*, 65: 382–388.

Davoine, C., Le Deunff, E., Ledger, N., Avice, J-C. and Billard, J-P. (2001). "Specific and constitutive expression of oxalate oxidase during the ageing of leaf sheaths of ryegrass stubble." *Plant Cell & Environment,* 24:1033–1043.

De Bold, A. J. (1985). "Atrial Natriuretic factor: a hormone produced by the heart." *Science,* 230: 767.

De Bold, A.J., Borenstein, B.H., Veress, A.T. and Sonnenberg, H. (1981). "A rapid and potent natriuretic response to intravenous injection of atrial myocardial extract in rats." *Life Sciences,* 28: 89 – 94.

Delaney, T. P. (1997). "Genetic dissection of acquired resistance to disease." *Plant Physiology,* 113: 5-12.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M. and Diaz-Orejas, R. (1998). "Replication and control of circular bacterial plasmids." *Microbiology & Molecular Biology Reviews,* 62: 434–464.

De Plater, G. M., Martin, R. L. and Milburn, P. J. (1998). "The natriuretic peptide (ovCNP-39) from platypus (*Ornithorhynchus anatinus*) venom relaxes the isolated uterus and promotes oedema and mast cell histamine release." *Toxicon,* 36: 847-857.

Donaldson, P. A., Anderson, T., Lane, B. G., Davidson, A. L. and Simmonds, D. H. (2001). "Soybean plants expressing an active oligomeric oxalate oxidase from the wheat *gf-2.8* (germin) gene are resistant to the oxalate-secreting pathogen *Sclerotina sclerotiorum*." *Physiological & Molecular Plant Pathology,* 59:297–307.

Dubendorff, J. W. & Studier, F. W. (1991). "Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor." *Journal of Molecular Biology,* 219: 45–59.

Durner, J., Wendehenne, D. and Klessig, D. F. (1998). "Defense gene induction in tobacco by nitric oxide, cyclic GMP and cyclic ADP-ribose" *Proceedings of the National Academy of Science USA,* 95: 10328–10333.

Edreva, E. (2005). "Pathogenesis-related proteins: Research progress in the last 15 years." *General & Applied Plant Physiology,* 31: 105-124.

Emanuelsson, O., Brunak, S., von Heijne, G. and Nielsen, H. (2007). "Locating proteins in the cell using TargetP, SignalP, and related tools." *Nature Protocols,* 2: 953-971.

Epstein, E. & Ludwig-Müller, J. (1993). "Indole-3-butyric acid in plants: occurrence, synthesis, metabolism and transport." *Physiologia Plantarum,* 88: 382-387.

Epstein, E. Baldi, B. G. and Cohen, J. D. (1984). "Chromatographic evidence for the presence of indole-3-acetyl-glutamate in seed of soybean." *Proceedings of Plant Growth Regulators Society of America*, 11: 184-188.

Epstein, E., Cohen, J. D. and Bandurski, R. S. (1980). "Concentration and metabolic turnover of indoles in germinating kernels of *Zea mays* L." *Plant Physiology,* 65: 415–421.

Faktor, O. & Raviv, M. (1997). "Inhibition of molt in *Spodoptera littoralis* larvae treated with soybean Bowman-Birk protease inhibitor." *Entomologia Expis Applicata,* 82: 109-113.

Farabaugh, P. J. (1978). "Sequence of the lac I gene." *Nature*, 274: 765.

Farrell, A. P. & Olson, K. R. (2000). "Cardiac natriuretic peptides: A physiological lineage of cardioprotective hormone?" *Physiological & Biochemical Zoology*, 73: 1-11.

Feller, S. M., Gagelmann, M. and Forssman, W. G. (1989). "Urodilatin: A newly described member of the ANP family." *Trends in Pharmacological Sciences,* 10: 93-94.

Ferry, N., Jouanin, L., Ceci, L. R., Mulligan, A., Emami, K., Gatehouse, J. A. and Gatehouse, A. M. (2005). "Impact of oilseed rape expressing the insecticidal serine protease inhibitor, mustard trypsin inhibitor-2 on the beneficial predator *Pterostichus madidus.*" *Molecular Ecology,* 14: 337-349.

Fickett, J. W. & Hatzigeorgiou, A. G. (1997). "Eukaryotic promoter prediction." *Genome Research,* 7: 861-878.

Fiers, M., Hause, G., Boutilier, K., Casamitjana-Martinez, E., Weijers, D., Offringa, R., van der Geest, L., van Lookeren Campagne, M. and Liu, C. (2004). "Mis-expression of the CLV3/ESR-like gene CLE19 in *Arabidopsis* leads to a consumption of root meristem." *Gene,* 327: 37-49.

Finkelstein, R. R., Gampala, S. S. L. and Rock, C. D. (2002). "Abscisic acid signalling in seeds and seedlings." *The Plant Cell,* 14: S15–S45.

Fletcher, J. C., Brand, U., Running, M. P., Simon, R., and Meyerowitz, E. M. (1999). "Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems." *Science,* 283: 1911–1914.

Fowkes, R. C. & McArdle, C. A. (2000). "C-type natriuretic peptide: An important neuroendocrine regulator?" *Trends in Endocrinology and Metabolism,* 11(8): 333-338.

Fujioka, S. (1999**).** "Natural occurrence of brassinosteroids in the plant kingdom." *In* A Sakurai, T Yokota, SD Clouse, eds, Brassinosteroids: Steroidal Plant Hormones. Springer-Verlag, Tokyo, pp 21-45.

Garbers, D. L. & Lowe, D. G. (1994). "Guanylyl cyclase receptors." *Journal of Biological Chemistry,* 269: 30741–30744.

Garthwaite, J., Southam, E., Boulton, C. L., Nielsen, E. B., Schmidt, K. and Mayer, B. (1995). "Potent and selective inhibition of nitric oxide-sensitive guanylyl cyclase by 1H-[1,2,4]oxadiazolo[4,3-α]quinoxalin-1-one." *Molecular Pharmacology,* 48: 184–188.

Gaudin, V., Vrain, T. and Jouanin, L. (1994). "Bacterial genes modifying hormonal balances in plants" *Plant Physiology & Biochemistry,* 32: 11–29.

Gehring, C. A. (1999). "Natriuretic peptides - A new class of plant hormone?" *Annals of Botany,* 83: 329-334.

Gehring, C. A. & Irving, H. R. (2003). "Natriuretic peptides – a class of heterologous molecules in plants." *The International Journal of Biochemistry & Cell Biology,* 1482: 1-5.

Gehring, C. A., Khalid, K. M., Toop, T. and Donald, J. A. (1996). "Rat natriuretic peptide binds specifically to plant membranes and induces stomatal opening." *Biochemical and Biophysical Research Communications*, 228: 739-744.

Gibson, L. J. (1994). "Pseudogenes and Origins." *Origins*, 21(2): 91-108.

Grace, M. L., Chandrasekharan, M. B., Hall, T. C. and Crowe, A. J. (2004). "Sequence and spacing of TATA box elements are critical for accurate initiation from the beta-phaseolin promoter." *Journal of Biological Chemistry*, 279: 8102-8110.

Gressent, F., Rahioui, I. and Rahbe, Y. (2003). "Characterization of a high-affinity binding site for the pea albumin 1b enterotoxin in the weevil *Sitophilus*." *European Journal of Biochemistry*, 270: 2429-2435.

Griffith, M. & Yaish, M. W. (2004). "Antifreeze proteins in overwintering plants: a tale of two activities." *Trends in Plant Science*, 9: 399–405.

Griffiths, A. J. F., Wessler, S. R., Lewontin, R. C. and Carroll, S. B. (2008). "Introduction to genetic analysis." 9th Edition, W. H. Freeman & Company, New York.

Grove, M. D., Spencer, F. G., Rohwedder, W. K., Mandava, N. B., Worley, J. F., Warthen Jr., J. D. Steffens, G. L. Flippen-Anderson J. L. and Carter-Cook J. Jr. (1979). "Brassinolide, a plant growth promoting steroid isolated from *Brassica napus* L. Pollen." *Nature*, 281: 216–217.

Grunwald, I., Rupprecht, I., Schuster, G. and Kloppstech, K. (2003). "Identification of guttation fluid proteins: the presence of pathogenesis-related proteins in non-infected barley plants." *Physiologia Plantarum*, 119: 192–202.

Gulyyaev, A. P. & Roussis, A. (2007). "Identification of conserved secondary structures and expansion segments in enod40 RNAs reveals new enod40 homologues in plants." *Nucleic Acids Research*, 35: 3144-3152.

Guzzo, F., Portaluppi, P., Grisi, R., Barone, S., Zampieri, S., Franssen, H. and Levi, M. (2005). "Reduction of cell size induced by enod40 in *Arabidopsis thaliana*." *Journal of Experimental Botany*, 56: 507-513.

Halliwell, C. M., Morgan, G., Ou, C. P. and Cass, A. E. (2001). "Introduction of a (poly)histidine tag in L-lactate dehydrogenase produces a mixture of active and inactive molecules." *Analytical Biochemistry*, 295: 257–261.

Hanai, H., Matsuno, T., Yamamoto, M., Matsubayashi, Y., Kamada, H. and Sakagami, Y. (2000). "A secreted peptide growth factor, phytosulfokine, acting as a stimulatory factor of carrot somatic embryo formation." *Plant Cell Physiology*, 41: 27-32.

Hannig, G. & Makrides, S. C. (1998). "Strategies for optimizing heterologous protein expression in *Escherichia coli*." *Trends in Biotechnology*, 16: 54–60.

Hefti, M. H., Caroline, J. G., der Toorn, V. V., Dixon, R. and Vervoort, J. (2001). "A novel purification method for histidine-tagged proteins containing a thrombin cleavage site." *Analytical Biochemistry*, 295: 180–185.

Heidecker, G. & Messing, J. (1986). "Structural analysis of plant genes." *Annual Review of Plant Physiology*, 37: 439-466.

Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999). "Plant cis-acting regulatory DNA elements (PLACE) database:1999." *Nucleic Acids Research*, 27(1): 297-300.

Hisamatsu, T., Koshioka, M., Kubota, S., Nishijima, T., Yamane, H., King, R. W. and Mander, L. N. (1998). "Isolation and identification of GA_{112} (12β-hydroxy-GA_{12}) in *Matthiola incana*." *Phytochemistry*, 47(1): 3-6.

Hobe, M., Muller, R., Grunewald, M., Brand, U. and Simon, R. (2003). "Loss of CLE40, a protein functionally equivalent to the stem cell restricting signal CLV3, enhances root waving in *Arabidopsis*." *Development Genes & Evolution*, 213: 371-381.

Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. and Stüber, D. (1988). "Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent." *Bio/Technology,* 6: 1321–1325.

Hochuli, E., Döbeli, H. and Schacher, A. (1987). "New metal chelate adsorbent selective for proteins and peptide containing neighbouring histidine residues." *Journal of Chromatography,* 411: 177–184.

Hoegen, E., Strömberg, A., Pihlgren, U. and Kombrink, E. (2002). "Primary structure and tissue-specific expression of the pathogenesis-related protein PR-1b in potato." *Molecular Plant Pathology,* 3: 329–345.

Hoffmann-Sommergruber, K. (2000). "Plant allergens and pathogenesis-related Proteins - What do they have in common?" *International Archives of Allergy & Immunology,* 122: 155–166.

Hon, W-C., Griffith, M., Mlynarz, A., Kwok, Y. C. and Yang, D. S. C. (1995). "Antifreeze proteins in winter rye are similar to pathogenesis-related proteins." *Plant Physiology,* 109: 879–889.

Horvath, B., Heidstra, R., Lados, M., Moerman, M., Spaink, H. P., Promé, J-C., van Kammen, A. and Bisseling, T. (1993**).** "Lipo-oligosaccharides of *Rhizobium* induce infection-related early nodulin gene expression in pea root hairs." *The Plant Journal,* 4(4): 727-733.

Huffaker, A. & Ryan, C. A. (2007). "Endogenous peptide defense signals in *Arabidopsis* differentially amplify signalling for the innate immune response." *Proceedings of the National Academy of Sciences USA*, 104: 10732-10736.

Huffaker, A., Pearce, G. and Ryan, C. A. (2006). "An endogenous peptide signal in *Arabidopsis* activates components of the innate immune response." *Proceedings of the National Academy of Sciences USA*, 103: 10098-10103.

Humphreys, M. H. & Valentin, J-P. (2000). "Natriuretic humoral agents." in Seldin DW, Giebisch G (eds): The Kidney: Physiology and Pathophysiology (ed 3). Philadelphia, PA, LippincottWilliams &Wilkins, pp 1371-1410.

Inagami, T., Misono, K. S., Fukumi, H., Maki, M., Tanaka, I., Takayanagi, R., Imada, T., Grammer, R. T., Naruse, M., Naruse, K., Pandey, K. N, Parmentier, M., Yasujima, M. and Abe, K. (1987). "Structure and physiological actions of rat atrial natriuretic factor." *Hypertension 10 (suppl I):* I113–I117.

Ito, Y., Nakanomyo, I., Motose, H., Iwamoto, K., Sawa, S., Dohmae, N. and Fukuda, H. (2006). "Dodeca-CLE peptides as suppressors of plant stem cell differentiation." *Science,* 313: 842-845.

Itoh, H. & Nakao, K. (1994). "Antagonism between the vascular renin-angiotensin and natriuretic peptide systems in vascular remodeling." *Blood Pressure 5 (suppl.),* 49-53.

Jackson, A. O. & Tailor, C. B. (1996). "Plant-microbe interactions: life and death at the interface." *Plant Cell Journal,* 8: 10.

Jacq, C., Miller, J. R. and Brownlee, G. G. (1977). "A pseudogene structure in 5S DNA of *Xenopus laevis*. *Cell,* 12: 109-120.

Jacquet, A., Daminet, V., Haumont, M., Garcia, L., Chaudoir, S., Bollen, A. and Biemans, R. (1999). "Expression of a recombinant *Toxoplasma gondii* ROP2 fragment as a fusion protein in bacteria circumvents insolubility and proteolytic degradation." *Protein Expression & Purification,* 17: 392–400.

Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R., Nordheim, A. and Stunnenberg, H. G. (1991). "Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccina virus." *Proceedings of the National Academy of Sciences USA,* 88: 8972–8976.

Jinn, T. L., Stone, J. M. and Walker, J. C. (2000). "HAESA, an *Arabidopsis* leucine-rich repeat receptor kinase, control floral organ abscission." *Genes & Development,* 4: 108-117.

Joffy, S. & Rosner, M. H. (2005). "Natriuretic peptides in ESRD." *American Journal of Kidney Diseases,* 46(1): 1-10.

Johansson, O., Alkema, W., Wasserman, W. W. and Lagergren, J. (2003). "Identification of functional clusters of transcription factor binding motifs in genome sequences: the MSCAN algorithm." *Bioinformatics,* 19: i169-i176.

John, P. C. L., Zhang, K., Dong, C., Diederich, L. and Wightman F. (1993). "$p34^{cdc2}$ related proteins in control of cell cycle progression, the switch between division and differentiation in tissue development, and stimulation of division by auxin and cytokinin." *Australian Journal of Plant Physiology,* 20: 503–526.

Jonasson, P., Liljeqvist, S., Nygren, P. A. and Stahl, S. (2002). "Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli*." *Biotechnology & Applied Biochemistry,* 35: 91–105.

Jones, D. T. (1999). "Protein secondary structure prediction based on position specific scoring matrices." *Journal of Molecular Biology,* 292: 195–205.

Jord´a, L., Conejero, V. and Vera, P. (2000). "Characterization of *P69E* and *P69F*, two differentially regulated genes encoding new members of the subtilisin-like proteinase family from tomato plants." *Plant Physiology,* 122: 67–73.

Kashiwagi, M., Katafuchi, T. and Kato, A. (1995). "Cloning and properties of a novel natriuretic peptide receptor NPR-D." *European Journal of Biochemistry,* 233: 102-109.

Katafuchi, T., Takashima, A., Kashiwagi, M., Hagiwara, H., Takei, Y. and Hirose, S. (1994). "Cloning and expression of eel natriuretic peptide receptor B (NPR-B) and its comparison with the mammalian counterparts." *European Journal of Biochemistry,* 222: 835-842.

Kecskemeti, V., Pacher, P., Pankucsi, C. and Nanasi, P. (1996). "Comparative study of cardiac electrophysiological effects of atrial natriuretic peptide." *Molecular and Cellular Biochemistry*, 161: 53-59.

Kende, H. & Zeevart, J. A. D. (1997). "The five 'classical' hormones." *The Plant Cell*, 9: 1197–1210.

Kende, H., Bradford, K. J., Brummell, D. A., Cho, H-T. and Cosgrove, D. J. (2004). "Nomenclature for members of the expansin superfamily of genes and proteins." *Plant Molecular Biology*, 55: 311–314.

Kehr, J., Buhtz, A. and Giavalisco, P. (2005). "Analysis of xylem sap proteins from *Brassica napus*." *BMC Plant Biology*, 5:11.

Kim, T. G., Kim, H. M., Lee, H. J., Shin, Y. J., Kwon, T. H., Lee, N. J., Jang, Y. S. and Yang, M. S. (2007). "Reduced protease activity in transformed rice cell suspension cultures expressing a proteinase inhibitor." *Protein Expression & Purification*, 53: 27-274.

Khripach, V., Zhabinskii, V. and deGroot, A. (2000). "Twenty years of brassinosteroids: Steroidal plant hormones warrant better crops for the XXI century." *Annals of Botany*, 86: 441–447.

Knox, B. & Suphioglu, C. (1996). "Environmental and molecular biology of pollen allergens." *Trends in Plant Science*, 1: 156–164.

Kobayashi, T., Eun, C.-H., Hanai, H., Matsubayashi, Y., Sakagami, Y. and Kamada, H. (1999). "Phytosulfokine-α, a peptidyl plant growth factor, stimulates cell division that leads to somatic embryogenesis in carrot." *Journal of Experimental Botany*, 50:1123-1128.

Koiwa, H., Bressan, R. A. and Hasegawa, P. M. (1997). "Regulation of protease inhibitors and plant defense." *Trends in Plant Science, 2*: 379-384.

Koller, K. J., Lowe, D. G., Bennett, G. L., Minamino, N., Kangawa, K., Matsuo, H. and Goeddel, D. V. (1991). "Selective activation of the B natriuretic peptide receptor by C-type natriuretic peptide (CNP)." *Science*, 252: 120-123.

Komarnytsky, S., Borisjuk, N., Yakoby, N., Garvey, A. and Raskin, I. (2006). "Cosecretion of protease inhibitor stabilizes antibodies produced by plant roots." *Plant Physiology*, 141: 1185-1193.

Kombrink, E. & Somssich, I. E. (1997). "Pathogenesis-related proteins and plant defense." In *The Mycota V, Part A. Plant Relationships*, ed. G Carroll, P Tudzynski, pp. 107–28. Berlin: Springer-Verlag.

Kondo, T., Sawa, S., Kinoshita, A., Mizuno, S., Kakimoto, T., Fukuda, H. and Sakagami, Y. (2006). "A plant peptide encoded by CLV3 identified by *in situ* MALDI-TOF MS analysis." *Science*, 313: 845-848.

Kone, B. C. (2001). "Molecular biology of natriuretic peptides and nitric oxide synthases." *Cardiovascular Research*, 51: 429-441.

Kourie, J. I. & Rive, M. J. (1999). "Role of natriuretic peptides in ion transport mechanisms." *Medicinal Research Reviews*, 19: 75–94.

Koyama, T., Okada, T., Kitajima, S., Ohme-Takagi, M., Shinshi, H. and Sato, F. (2003). "Isolation of tobacco ubiquitin-conjugating enzyme cDNA in a yeast two-hybrid system with tobacco ERF3 as bait and its characterization of specific interaction." *Journal of Experimental Botany*, 54: 1175-1181.

Lamb, C. & Dixon, R. A. (1997). "The oxidative burst in plant disease resistance." *Annual Reviews of Plant Physiology and Plant Molecular Biology*, 48: 251-275.

Lang, R. E., Tholken, H., Ganten, D., Luft, F. C., Ruskoaho, H. and Uger, T. (1985). "Atrial Natriuretic factor-a circulating hormone stimulated by volume loading" *Nature (London)*, 314: 264.

Lang, J. M., Eisinger, W. R. and Green, P. B. (1982). "Effects of ethylene on the orientation of microtubules and cellulose microfibrils of pea epicotyl cells with polylamellate cell walls." *Protoplasma,* 110: 5–14.

Lareau, L. F., Green, R. E., Bhatnagar, R. S. and Brenner, S. E. (2004). "The evolving roles of alternative splicing." *Current Opinion in Structural Biology,* 14: 273-282.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. and Higgins, D. G. (2007). "ClustalW2 and ClustalX version 2." *Bioinformatics,* 23(21): 2947-2948.

LaRosa, P. C., Chen, Z., Nelson, D. E., Singh, N. K., Hasegawa, P. M., Bressan, R. A. (1992). "Osmotin gene expression is posttranscriptionally regulated." *Plant Physiology,* 100: 409–415.

Laursen, B. S., Steffensen, S., Hedegaard, J., Moreno, J. M., Mortensen, K. K. and Sperling-Petersen, H. U. (2002). "Structural requirements of the mRNA for intracistronic translation initiation of the enterobacterial *infB* gene." *Genes Cells,* 7: 901–910.

Lay, F. T. & Anderson, M. A. (2005). "Defensins—components of the innate immune system in plants." *Current Protein & Peptide Science,* 6: 85–101.

Lee, Y. & Kende, H. (2002). "Expression of α-expansin and expansin-like genes in deepwater rice." *Plant Physiology,* 130: 1396-1405.

Lee, Y., Choi, D. and Kende, H. (2001). "Expansins: everexpanding numbers and functions." *Current Opinion in Plant Biology,* 4: 527–532.

Leon, P., & Sheen, J. (2003). "Sugar and hormone connections." *TRENDS in Plant Science,* 8(3): 110–116.

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P. and Rombauts, S. (2002). "PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences." *Nucleic Acids Research,* 30(1): 325-327.

Lesk, A. M. (2005). "Introduction to bioinformatics." 2nd Edition, Oxford University Press Inc., New York.

Levin, E. R. (1993). "Natriuretic peptide C-receptor: more than a clearance receptor." *American Journal of Physiology,* 264: E483-E489.

Li, Y., Jones, L. and McQueen-Mason, S. (2003). "Expansins and cell growth." *Current Opinion in Plant Biology,* 6: 603–610.

Li, Y., Darley, C. P., Ongaro, V., Fleming, A., Schipper, O., Baldauf, S. L. and McQueen-Mason, S. J. (2002). "Plant expansions are a complex multigene family with an ancient evolutionary origin." *Plant Physiology,* 128: 854–864.

Linder, M. & Teeri, T. T. (1997). "The roles and function of cellulose-binding domains." *Journal of Biotechnology*, 57: 15– 28.

Lindsey, K., Casson, S. and Chilley, P. (2002). "Peptides: new signalling molecules in plants." *TRENDS in Plant Science,* 7(2): 78–83.

Long, M. & Langley, C. H. (1993). "Natural selection and the orgin of jingwei, a chimeric processed functional gene in *Drosophila.*" *Science,* 260: 91-95.

Ludidi, N. & Gehring, C. (2003). "Identification of a novel protein with guanylyl cyclase activity in *Arabidopsis thaliana*." *Journal of Biological Chemistry,* 278(8): 6490-6494.

Ludidi, N., Morse, M., Sayed, M., Wherret, T., Shabala, S. and Gehring, C. (2004). "A recombinant plant natriuretic peptide causes rapid and spatially differentiated K^+, Na^+ and H^+ flux changes in *Arabidopsis thaliana* roots." *Plant Cell Physiology,* 45: 1093–1098.

Ludidi, N. N., Heazlewood, J. L., Seoighe, C. J., Irving, H. R. and Gehring, C. A. (2002). "Expansin-like molecules: Novel functions derived from common domains." *Journal of Molecular Evolution*, 54: 587-594.

MacMillan, J. (1997). "Biosynthesis of the gibberellin plant hormones." *Natural Product Reports*, 14: 221–243.

Martin, D. R., Pevahouse, J. B., Trigg, D. J., Vesely, D. L. and Buerkert, J. E. (1990). "Three peptides from the ANF prohormone NH_2-terminus are natriuretic and/or kaliuretic." *American Journal of Physiology*, 258: F1401–1408.

Maryani, M. M., Morse, M. V., Bradley, G., Irving, H. R., Cahill, D. M. and Gehring, C. A. (2003). "*In situ* localization associates biologically active plant natriuretic peptide immuno-analogues with conductive tissue and stomata." *Journal of Experimental Botany*, 54: 1553–1564.

Maryani, M. M., Bradley, G., Cahill, D. M. and Gehring, C. A. (2001). "Natriuretic peptides and immunoreactants modify osmoticum-dependent volume changes in *Solanum tuberosum* L. mesophyll cell protoplasts." *Plant Science*, 161: 443–452.

Maryani, M. M., Shabala, S. N. and Gehring, C. A. (2000). "Plant natriuretic peptide immunoreactants modulate plasma-membrane H^+ gradients in *Solanum tuberosum L*. leaf tissue vesicles." *Archives of Biochemistry & Biophysics*, 376: 456-458.

Matsubayashi, Y. & Sakagami, Y. (2006). "Peptide hormones in plants." *Annual Review of Plant Biology*, 57: 649-674.

Matsubayashi, Y. & Sakagami, Y. (1996). "Phytosulfokine, sulphated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L." *Proceedings of the National Academy of Sciences USA*, 93: 7623–7627.

Matsubayashi, Y., Yang, H. and Sakagami, Y. (2001). "Peptide signals and their receptors in higher plants." *Trends in Plant Science*, 6(12): 573-577.

Matsubayashi, Y., Takagi, L., Omura, N., Morita, A. and Sakagami, Y. (1999). "The endogenous sulfated pentapeptide, phytosulfokine-α, stimulates tracheary element differentiation of isolated mesophyll cells of *Zinnia elegans.*" *Plant Physiology,* 120:1043-1048.

Matsubayashi, Y., Takagi, L., and Sakagami, Y. (1997). "Phytosulfokine- α, a sulfated pentapeptide, stimulates the proliferation of rice cells by means of specific high- and low-affinity binding sites." *Proceedings of the National Academy of Sciences USA*, 94: 13357–13362.

Matthews, C. K., van Holde, K. E. and Ahern, K. G. (2000). "Biochemistry." 3rd Edition, Addison Wesley Longman, Inc. Benjamin/Cummings, CA.

McDonald, K. L. & Cahill, D. M. (1999). "Influence of abscisic acid and the abscisic acid biosynthesis inhibitor, Norflurazon, on interactions between *Phytophthora sojae* and soybean (*Glycine max*)." *European Journal of Plant Pathology,* 105: 651-658.

McDougall, G. J. & Fry, S. C. (1990). "Xyloglucan oligosaccharides promote growth and activate cellulase: Evidence for a role of cellulase in cell expansion." *Plant Physiology*, 93: 1042-1048.

McDougall, G. J. & Fry, S. C. (1989). "Anti-auxin activity of xyloglucan oligosaccharides: The role of groups other than the terminal α-L-fucose residue." *Journal of Experimental Botany,* 40: 233-239.

McGaw, B. A. & Burch, L. R. (1995). "Cytokinin biosynthesis and metabolism." In Davies, P. J. (ed). Plant Hormones – Physiology, Biochemistry and Molecular Biology. Kluwer Academic. Netherlands. pp. 98-117.

McGurl, B., Pearce, G., Orozco-Cardenas, M. and Ryan, C.A. (1992). "Structure, expression, and antisense inhibition of the systemin precursor gene." *Science,* 255(5051): 1570–1573.

McKeon, T. A., Fernandez-Maculet, J. A. and Yang, S. F. (1995). "Biosynthesis and metabolism of ethylene." In Davies, P. J. (ed). Plant Hormones – Physiology, Biochemistry and Molecular Biology. Kluwer Academic. Netherlands. pp. 118-139.

McQueen-Mason, S. J. & Cosgrove, D. J. (1995). "Expansin mode of action on cell walls. Analysis of wall hydrolysis, stress relaxation, and binding." *Plant Physiology*, 107(1): 87–100.

McQueen-Mason, S. J., Durachko, D. M. and Cosgrove, D. J. (1992). "Two endogenous proteins that induce cell wall extension in plants." *Plant Cell*, 4: 1425-1433.

Meier, S., Bastian, R., Donaldson, L., Murray, S., Bajic, V. and Gehring, C. (2008). "Co-expression and promoter content analyses assign a role in biotic and abiotic stress responses to plant natriuretic peptides." *BMC Plant Biology*, 8: 24-36.

Meindl, T., Boller, T. and Felix, G. (1998). "The plant wound hormone systemin binds with the N-terminal part to its receptor but needs the C-terminal part to activate it." *Plant Cell*, 10: 1561–1570.

Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A. and Kondorosi, E. (2003). "A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs." *Plant Physiology*, 132:161-173.

Metraux, J. –P. (2002). "Recent breakthroughs in the study of salicylic acid biosynthesis." *TRENDS in Plant Science*, 7(8): 332-334.

Meyerowitz, E. M. (1999). "Plant, animals and the logic of development." *Trends in Cell Biology*, 9: M65 – M68.

Michener, M. L., Gierse, J. K., Seetharam, R., Fok, K. F., Olins, P. O., Mai, M. S. and Needleman, P. (1986). "Proteolytic processing of atriopeptin prohormone." *Molecular Pharmacology*, 30(6): 552–557.

Minami, E., Kouchi, H., Cohn, J. R., Ogawa, T. and Stacey, G. (1996). "Expression of the early nodulin, ENOD40, in soybean roots in response to various lipo-chitin signal molecules." *The Plant Journal*, 10(1): 23-32.

Mishima, M., Takayama, S., Sasaki, K.-I., Jee, J.-G., Kojima, C., Isogai, C. and Shirakawa, M. (2003). "Structure of the male determinant factor for *Brassica* self-incompatibility." *Journal of Biology Chemistry*, 278: 36389-36395.

Mohr, P. G. & Cahill, D. M. (2003). "Abscisic acid influences the susceptibility of *Arabidopsis thaliana* to *Pseudomonas syringae* pv. tomato and *Peronospora parasitica*." *Functional Plant Biology*, 30: 461-469.

Monteiro, A. M., Crozier, A. and Sandberg, G. (1988). "The biosynthesis and conjugation of indole-3-acetic acid in germinating seed and seedlings of *Dalbergia dolchipetala.*" *Planta*, 174: 561–568.

Morishita, Y., Sano, T. and Ando, K. (1991). "Microbial polysaccharide, HS-142-1 competitively and selectively inhibits ANP binding to its guanylyl cyclase containing receptor." *Biochemical & Biophysical Research Communications*, 176: 949-957.

Morse, M., Pironcheva, G. and Gehring, C. (2004). "*AtPNP-A* is a systemically mobile natriuretic peptide immunoanalogue with a role in *Arabidopsis thaliana* cell volume regulation." *FEBS Letters*, 556: 99 – 103.

Mülsch, A., Busse, R., Liebau, S. and Forstermann, U. (1988). "LY 83583 interferes with the release of endothelium-derived relaxing factor and inhibit soluble guanylate cyclase." *Journal of Pharmacology and Experimental Therapeutics*, 247: 283–288.

Mussig, C. & Altman, T. (1999). "Physiology and molecular mode of action of brassinosteroids." *Plant Physiology & Biochemistry*, 37(5): 363–372.

Nakamura, M., Tsunoda, T. and Obokata, J. (2002). "Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator." *Plant Journal*, 29: 1-10.

Nakashima, K., Fujita, Y., Katsura, K., Maruyama, K., Narusaka, Y., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2006). "Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of *Arabidopsis*." *Plant Molecular Biology*, 60: 51-68.

Narita, N. N., Moore, S., Horiguchi, G., Kubo, M., Demura, T., Fukuda, H., Goodrich, J. and Tsukaya, H. (2004). "Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in *Arabidopsis thaliana*." *Plant Journal*, 38: 699-713.

Narváez-Vásquez, J. & Ryan, C. (2004). "The cellular localization of prosystemin: a functional role for phloem parenchyma in systemic wound signalling." *Planta*, 218: 360-369.

Narváez-Vásquez, J., Pearce, G. and Ryan, C. A. (2005). "The plant cell wall matrix harbours a precursor of defense signalling peptides." *Proceedings of the National Academy of Sciences USA*, 102: 12974 – 12977.

Neale, A. D., Wahleithner, J. A., Lund, M., Bonnett, H. T., Kelly, A., Meeks-Wagner, D. R., Peacock, W. J. and Dennis, E. S. (1990). "Chitinase, β-1,3-glucanase, osmotin, and extension are expressed in tobacco explants during flower formation." *The Plant Cell*, 2: 673-684.

Needleman, P., Blaine, E. H., Greenwald, J. E., Michener, M. L., Saper, C. B., Stockmann, P.T. and Tolunay, H. E. (1989). "The biochemical pharmacology of atrial peptides." *Annual Review of Pharmacology & Toxicology*, 29: 25–54.

Nelson, D. L. & Cox, M. M. (2005). "Lehninger: Principles of biochemistry." 4th Edition, W. H. Freeman and Company, New York.

Nembaware, V., Seoighe, C., Sayed, M. and Gehring, C. (2004). "A plant natriuretic peptide-like gene in the bacterial *Xanthomonas axonopodis* may induce hyper-hydration in the plant host: a hypothesis of molecular mimicry." *BMC Evolutionary Biology*, 4: 10-17.

Newbury, S. F., Smith, N. H., Robinson, E. C., Hiles, I. D. and Higgins, C. F. (1987). "Stabilization of translationally active mRNA by prokaryotic REP sequences." *Cell,* 48: 297–310.

Newton, R. P., Roef, L., Witters, E. and Van Onckelen, H. (1999). "Tansley Review No. 106 - Cyclic nucleotides in higher plants: the enduring paradox." *New Phytologist,* 143: 427-455.

Ni, J. & Clark, S. E. (2006). "Evidence for functional conservation, sufficiency, and proteolytic processing of the CLAVATA3 CLE domain." *Plant Physiology,* 140: 726-733.

Nikolov, D. B. & Burley, S. K. (1997). "RNA polymerase II transcription initiation: a structural view." *Proceedings of the National Academy of Sciences USA,* 94: 15-22.

Normanly, J. (1997). "Auxin metabolism." *Physiologia Plantarum,* 100: 431–442.

Normanly, J., Slovin, J. P. and Cohen, J. D. (1995). "Rethinking auxin biosynthesis and metabolism." *Plant Physiology,* 107: 323-329.

Oh, M. H. & Clouse, S. D. (1998). "Brassinolide affects the rate of cell division in isolated leaf protoplasts of *Petunia hybrida*." *Plant Cell Reports,* 17: 921-924.

Olesen, S-P., Drejer, J., Axelsson, O., Moldt, P., Bang, L., Nielsen-Kudsk, J. E., Busse, R. and M¨ulsch, A. (1998). "Characterization of NS 2028 as a specific inhibitor of soluble guanylyl cyclase." *British Journal of Pharmacology,* 123: 299–309.

Olsen, A. N., Mundy, J. and Skriver, K. (2002). "Peptomics, identification of novel cationic *Arabidopsis* peptides with conserved sequence motifs." *In Silico Biology,* 2: 441-451.

Olson, K. R. (1998). "Hormone metabolism by the fish gill." *Comparative Biochemistry & Physiology,* 119A: 55-65.

Orozco-Cardenas, M. & Ryan, C. A. (1999). "Hydrogen peroxide is generated systemically in plant leaves by wounding and systemin via the octadecanoid pathway." *Proceedings of the National Academy of Sciences USA*, 96: 6553-6557.

Park, C. –J., An, J. –M., Shin, Y. –C., Kim, K. –J., Lee, B. –J. and Paek, K. –H. (2004). "Molecular characterization of pepper germin-like protein as novel PR-16 family of pathogenesis-related proteins isolated during the resistance response to viral and bacterial infection." *Planta*, 219: 797-806.

Passardi, F., Penel, C. and Dunand, C. (2004). "Performing the paradoxical: How plant peroxidases modify the cell wall." *Trends in Plant Science*, 9: 534–540.

Pearce, G. & Ryan, C. A. (2003). "Systemic signaling in tomato plants for defense against herbivores. Isolation and characterization of three novel defense-signaling glycopeptide hormones coded in a single precursor gene." *Journal of Biological Chemistry*, 278: 30044-30050.

Pearce, G., Siems, W. F., Bhattacharya, R., Chen, Y. C. and Ryan, C. (2007). "Three hydroxyproline-rich glycopeptides derived from a single petunia polyprotein precursor activate defensin I, a pathogen defense response gene." *Journal of Biological Chemistry*, 282: 17777-17784.

Pearce, G., Moura, D.S., Stratmann, J., and Ryan, C.A. (2001). "RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development." *Proceedings of the National Academy of Sciences USA*, 98: 12843–12847.

Pearce, G., Strydom, D., Johnson, S. and Ryan, C.A. (1991). "A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins." *Science*, 253(5022): 895-898.

Pedersen, A. G., Baldi, P., Chauvin, Y. and Brunak, S. (1999). "The biology of eukaryotic promoter prediction – a review." *Computers & Chemistry*, 23: 191-207.

Penson, S. P., Schuurink, R. C., Fath, A., Gubler, F., Jacobsen, J. V. and Jones, R. L. (1996). "cGMP is required for gibberellic acid-induced gene expression in barley aleurone layer." *The Plant Cell*, 8: 2325–2333.

Percival, F. W. & Bandurski, R. S. (1976). "Esters of indole-3-acetic acid from *Avena* seeds." *Plant Physiology*, 58: 60-67.

Petrov, V., Amery, A. and Lijnen, P. (1994). "Role of cyclic GMP in atrial-natriuretic-peptide stimulation of erythrocyte Na^+/H^+ exchange." *European Journal of Biochemistry*, 221: 195-199.

Pharmawati, M., Maryani, M. M., Nikolakopoulos, T., Gehring, C. A. and Irving, H. R. (2001). "Cyclic GMP modulates stomatal opening induced by natriuretic peptides and immunoreactive analogues." *Plant Physiology & Biochemistry*, 39: 1-10.

Pharmawati, M., Shabala, S. N., Newman, I. A. and Gehring, C. A. (1999). "Natriuretic peptides and cGMP modulate K+, Na+ and H+ fluxes in *Zea mays* roots." *Molecular & Cell Biology Research Communications*, *2*: 53–57.

Pharmawati, M., Billington, T. and Gehring, C. A. (1998a). "Stomatal guard cell responses to kinetin and natriuretic peptides are cGMP dependent." *Cellular & Molecular Life Sciences*, 54: 272-276.

Pharmawati, M., Gehring, C. A. and Irving, H. R. (1998b). "An immunoaffinity purified plant natriuretic peptide analogue modulates cGMP levels in the *Zea mays* root stele." *Plant Science*, 137: 107-115.

Pidgeon, G. B., Richards, A. M., Nicholls, M. G., Espiner, E. A., Yandle, T. G. and Frampton, C. (1996). "Differing metabolism and bioactivity of atrial and brain natriuretic peptides in essential hypertension." *Hypertension*, 27(4): 906–913.

Pien, S., Wyrzykowska, J., McQueen-Mason, S. J., Smart, C. and Fleming, A. J. (2001). "Local expression of expansin induces the entire process of leaf development and modifies leaf shape." *Proceedings of the National Academy of Sciences USA*, 98: 11812–11817.

Pieterse, C. M. J., van Wees, S. C. M., Hoffland, E., van Pelt, J. A. and van Loon, L. C. (1996). "Systemic resistance in *Arabidopsis* induced by biocontrol bacteria is independent of salicylic acid accumulation and pathogenesis-related gene expression." *Plant Cell*, 8: 1225-1237.

Piskornik, R. & Bandurski, R. S. (1972). "Purification and partial characterization of a glucan containing indole-3-acetic acid." *Plant Physiology*, 50: 76-182.

Plesch, G., Ehrhardt, T. and Mueller-Roeber, B. (2001). "Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression." *Plant Journal*, 28: 455-464.

Poole, E. S., Brown, C. M. and Tate, W. P. (1995). "The identity of the base following the stop codon determines the efficiency of in vivo translational termination in *Escherichia coli*." *EMBO Journal*, 14: 151–158.

Porath, J., Carlsson, J., Olsson, I. and Belfrage, G. (1975). "Metal chelate affinity chromatography, a new approach to protein fractionation." *Nature*, 258: 598–599.

Quirino, B. F., Noh, Y-S., Himelblau, E. and Amasino, R. M. (2000). "Molecular aspects of leaf senescence." *Trends in Plant Science*, 5: 278–82.

Rademaker, M. T., Charles, C. J., Lewis, L. K., Yandle, T. G., Cooper, G. J. S., Coy, D. H., Richards, A. M. and Nicholls, M. G. (1997) "Beneficial hemodynamic and renal effects of adrenomedullin in an ovine model of heart failure." *Circulation*, 96**:** 1983–1990.

Rafudeen, S., Gxaba, G., Makgoke, G., Bradley, G., Pironcheva, G., Raitt, L., Irving, H. and Gehring, C. (2003). "A role for plant natriuretic peptide immune-analogues in NaCl- and drought-stress responses." *Physiologia Plantarum*, 119: 554–562.

Rance, I., Fournier, J. and Esquerre-Taugaye, M-T. (1998). "The incompatible interaction between Phytophthora parasitica var. nicotinae race 0 and tobacco is suppressed in transgenic plants expressing antisense lipoxygenase sequences." *Proceedings of the National Academy of Sciences U.S.A.*, 95: 6554-6559.

Rank, K. B., Mildner, A. M., Leone, J. W., Koeplinger, K. A., Chou, K. C., Tomasselli, A. G., Heinrikson, R. L. and Sharma, S. K. (2001). "[W206R]-procaspase 3: an inactivateable substrate for caspase 8." *Protein Expression & Purification*, 22: 258–266.

Raskin, I. (1995). "Salicylic acid." In Davies, P. J. (ed). Plant Hormones – Physiology, Biochemistry and Molecular Biology. Kluwer Academic. Netherlands. Pp. 188-205.

Rawat, R., Xu, Z. F., Yao, K. M. and Chye, M. L. (2005). "Identification of cis-elements for ethylene and circadian regulation of the *Solanum melongena* gene encoding cysteine proteinase." *Plant Molecular Biology*, 57: 629-643.

Rawlings, N. D., Morton, F. R. and Barrett, A. J. (2006). "MEROPS: the peptidase database." *Nucleic Acids Research*, 34: D270-D272.

Razem, F. A., Baron, K. and Hill, R. D. (2006). "Turning on gibberellin and abscisic acid signaling." *Current Opinion in Plant Biology*, 9: 454–459.

Reeck, G. R., Kramer, K. J., Barker, J. E., Kanost, M. R., Fabrick, J. A. and Behnke, C. A. (1997). "Proteinase inhibitors and resistance of transgenic plants to insects." In: *Advances in Insect Control: the Role of Transgenic Plants* (Carozzi, N. and Koziel M., eds), pp. 157-183. London: Taylor & Francis.

Rep, M., Dekker, H. L., Vossen, J. H, De Boer, A. D. and Houterman, P. M. (2003). "A tomato xylem sap protein represents a new family of small cysteine-rich proteins with structural similarity to lipid transfer proteins." *FEBS Letters*, 534: 82–86.

Rep, M., Dekker, H. L., Vossen, J. H., De Boer, A. D. and Houterman, P. M. (2002). "Mass spectrometric identification of isoforms of PR proteins in xylem sap of fungus-infected tomato." *Plant Physiology*, 130: 904–17.

Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G.D. and Gold, L. (1992). "Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site." *Molecular Microbiology*, 6: 1219–1229.

Rivard, D., Anguenot, R., Brunelle, F., Le, V. Q., Vèzina, L. P., Trèpanier, S. and Michaud, D. (2006). "An in-built proteinase inhibitor system for the protection of recombinant proteins recovered from transgenic plants." *Plant Biotechnology Journal*, 4: 359-368.

Röhrig, H., Schmidt, J., Miklashevichs, E., Schell, J. and John, M. (2002). "Soybean ENOD40 encodes two peptides that bind to sucrose synthase." *Proceedings of the National Academy of Sciences USA*, 99: 1915-1920.

Rosenzweig, A. & Seidman, C. E. (1991). "Atrial natriuretic factor and related peptide hormones." *Annual Review of Biochemistry*, 60: 229-255.

Ryan, C. A. (1992). "The search for the proteinase inhibitor-inducing factor, PIFF." *Plant Molecular Biology*, 19: 123-133.

Ryan, C. A. (1990). "Protease inhibitors in plants: genes for improving defenses against insects and pathogens." *Annual Review of Phytopathology*, *28*: 425-449.

Ryan, C. A., Huffaker, A. and Yamaguchi, Y. (2007). "New insights into innate immunity in Arabidopsis." *Cellular Microbiology*, 9: 1902–1908.

Sagili, R. R., Pankiw, T. and Zhu-Salzmman, K. (2005). "Effects of soybean trypsin inhibitor on hypopharyngeal gland protein content, total midgut protease activity and survival of the honey bee (*Apis mellifera L.*)" *Journal of Insect Physiology*, 51: 953-957.

Schaller, A. & Ryan, C. A. (1995). "Systemin - a polypeptide defense signal in plants." *Bioessays*, 18(1): 27-33.

Schmidt, M., Tuominen, N., Johansson, T., Weiss, S. A., Keinamen, K. and Oker-Blom, C. (1998). "Baculovirus-mediated large-scale expression and purification of a polyhistidine-tagged Rubrella virus capsid protein." *Protein Expression & Purification,* 12: 323–330.

Schopfer, C. R. & Nasrallah, J. B. (2000). "Self-incompatibility. Prospects for a novel putative peptide-signaling molecule." *Plant Physiology,* 124: 935–939.

Schopfer, C. R., Nasrallah, M. E. and Nasrallah, J. B. (1999). "The male determinant of self-incompatibility in *Brassica."* *Science,* 286: 1697–1700.

Schrammel, A., Behfrends, S., Schmidt, K., Koesling, D. and Mayer, B. (1996). "Characterization of 1H-[1,2,4]oxadiazolo[4,3-α]quinoxalin-1-one as a heme-site inhibitor of nitric oxide-sensitive guanylyl cyclase." *Molecular Pharmacology,* 50: 1–5.

Schroeder, J. I., Kwak, J. M. and Allen, G. J. (2001). "Guard cell abscisic acid signalling and engineering drought hardiness in plants." *Nature,* 410: 327–330.

Schulz, S., Singh, S., Bellet, R. A., Singh, G., Tubb, D. J., Chin, H. and Garbers, D. L. (1989). "The primary structure of a plasma membrane guanylate cyclase demonstrates diversity within this new receptor family." *Cell,* 58: 1155–1162.

Schwartz, D., Geller, D., Manning, P. T., Siegel, N. R., Fok, K. F., Smith, C. E. and Needleman, P. (1985). "Ser-Leu-Arg-Arg-atriopeptin III: the major circulating form of atrial peptide." *Science,* 229: 397-400.

Schwarz, E., Scherer, G., Hobom, G. and Kössel, H. (1978). "Nucleotide sequence of cro, cII and part of the O gene in phage lambda DNA." *Nature,* 272: 410-414.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003). "SWISS-MODEL: an automated protein homology-modeling server." *Nucleic Acids Research,* 31(13): 3381–3385.

Shah, J. & Klessig, D. F. (1999). "Salicylic acid: signal perception and transduction." In Biochemistry and Molecular Biology of Plant Hormones, Edited by Libbenga K, Hall M, Hooykaas, P. J. J., London: Elsevier, pp 513-541.

Shang, C., Shibahara, T., Hanada, K., Iwafune, Y. and Hirano, H. (2004). "Mass spectrometric analysis of posttranslational modifications of a carrot extracellular glycoprotein." *Biochemistry,* 43: 6281-6292.

Shapiro, L. & Harris, T. (2000). "Finding function through structural genomics." *Current Opinion in Biotechnology,* 11: 31-35.

Sharma, Y. J., Leon, J., Raskin, I. and Davis K. R. (1996). "Ozone-induced responses in Arabidopsis thaliana: The role of salicylic acid in the accumulation of defense related transcripts and induced resistance." *Proceedings of the National Academy of Sciences USA,* 93: 5099-5104.

Shimekake, Y., Ohta, Y. and Nagata, K. (1994). "C-type natriuretic peptide stimulates secretion of growth hormone from rat-pituitary-derived GH3 cells via a cyclic-GMP-mediated pathway." *European Journal of Biochemistry,* 222: 645–650.

Shirsat, A., Wilford, N., Croy, R.and Boulter, D. (1989). "Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco." *Molecular & General Genetics,* 215: 326-331.

Sigrist, C. J. A., Cerrutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A. and Bucher, P. (2002). "PROSITE: A documented database using patterns and profiles as motif descriptors." *Briefings in Bioinformatics,* 3(3): 265-274.

Simpson, S. D., Nakashima, K., Narusaka, Y., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2003). "Two different novel cis-acting elements of erd1, a clpA homologous *Arabidopsis* gene function in induction by dehydration stress and dark-induced senescence." *Plant Journal,* 33: 259-270.

Singh, N. K., LaRosa, P. C., Handa, A. K., Hasegawa, P. M. and Bressan, R. A. (1987). "Hormonal regulation of protein synthesis associated with salt tolerance in plant cells." *Proceedings of the National Academy of Sciences USA,* 84: 739-743.

Singh, N. K., Handa, A. K., Hasegawa, P. M. and Bressan, R. A. (1985). "Proteins associated with adaptation of cultured tobacco cells to NaCI." *Plant Physiology,* 79: 126- 137.

Solano, R., Nieto, C., Avila, J., Canas, L., Dias, I. and Paz-Ares, J. (1995). "Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB Ph3) from *Petunia hybrida*." *EMBO Journal,* 14: 1773-1784.

Sponsel, V. M. (1995). "Gibberellin biosynthesis and metabolism." In Davies, P. J. (ed) Plant Hormones – Physiology, Biochemistry and Molecular Biology. Kluwer Academic. Netherlands. pp. 66-97.

Staswick, P. E. (1995). "Jasmonate activity in plants." In Davies, P. J. (ed) Plant Hormones – Physiology, Biochemistry and Molecular Biology. Kluwer Academic. Netherlands. pp. 179-187.

Stearns, S. C. & Hoekstra, R. F. (2005). "Evolution: An introduction." 2nd Edition, Oxford University Press, USA.

Stenstrom, C. M., Jin, H., Major, L. L., Tate, W. P. and Isaksson, L. A. (2001). "Codon bias at the 3`-side of the initiation codon is correlated with translation initiation efficiency in *Escherichia coli*." *Gene,* 263: 273–284.

Stenvik, G.-E., Butenko, M. A., Urbanowicz, B. R., Rose, J. K. C. and Aalen, R. B. (2006). "Overexpression of Inflorescence Deficient in Abscission activates cell separation in vestigial abscission zones in *Arabidopsis.*" *Plant Cell,* 18: 1467-1476.

Stevens, R. C. (2000). "Design of high-throughput methods of protein production for structural biology." *Structure Fold Des.,* 8: R177–R185.

Sticher, L., Mauch-Mani, B. and Mètraux, J.-P. (1997). "Systemic acquired resistance." *Annual Review of Phytopathology*, 35: 235-270.

Stintzi, A., Heitz, T., Kauffmann, S., Legrand, M. and Fritig, B. (1991). "Identification of a basic pathogenesis-related thaumatin-like protein of virus-infected tobacco as osmotin." *Physiological & Molecular Plant Pathology,* 38: 137-146.

Stockand, J. D. & Sansom, S. C. (1996). "Role of large, Ca-activated K channels in regulation by nitroprusside and ANP of mesangial contraction." *American Journal of Physiology,* 270: C1773-C1779.

Studier, F. W. & Moffatt, B. A. (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." *Journal of Molecular Biology,* 189: 113-130.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990). "Use of T7 RNA polymerase to direct expression of cloned genes." *Methods in Enzymology,* 185: 60-89.

Stueber, D., Matile, H. and Garotta, G. (1990). "System for high-level production in Escherichia coli and rapid purification of recombinant proteins: application to epitope mapping, preparation of antibodies, and structure-function analysis." In: Immunological Methods, Lefkovits, I. and Pernis, B., eds., vol. IV, Academic Press, New York, pp. 121-152.

Sudoh, T., Kangawa, K., Minamino, N. and Matsuo, H. (1988). "A new natriuretic peptide in porcine brain." *Nature,* 332: 78-81.

Sudoh, T., Minamino, N., Kangawa, K. & Matsuo, H. (1990). "C-type natriuretic peptide (CNP): a new member of natriuretic peptide family identified in porcine brain." *Biochemical & Biophysical Research Communications,* 168:863-870.

Summers, D. (1998). "Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability." *Molecular Microbiology,* 29: 1137–1145.

Sutcliffe, J. G. (1979). "Complete nucleotide sequence of the *E. coli* plasmid pBR322." *Cold Spring Harbor Symposia on Quantitative Biology,* 43: 77-90.

Suwastika, I. N. & Gehring, C. A. (1998). "Natriuretic peptide hormones promote radial water movements from the xylem of *Tradescantia* shoots." *Cellular & Molecular Life Sciences,* 54: 1161-1167.

Suwastika, I. N., Toop, T., Irving, H. R. and Gehring, C. A. (2000). "*In situ* and *in vitro* binding of natriuretic peptide hormones in *Tradescantia multiflora*." *Plant Biology,* 2: 1-3.

Suzuki, T., Yamazaki, T. and Yazaki, Y. (2001). "The role of natriuretic peptides in the cardiovascular system." *Cardiovascular Research,* 51(3): 489-494.

Suzuki, E., Hirata, Y., Kohmoto, O., Sugimoto, T., Hayakawa, H., Matsuoka, H., Sugimoto, T., Kojima, M., Kangawa, K., Minamino, N. and Matso, H. (1992). "Cellular mechanisms for synthesis and secretion of atrial natriuretic peptide and brain natriuretic peptide in cultured rat atrial cells." *Circulation Research,* 71(5): 1039-1048.

Syvertsen, J. P., Bauscher, M. G. and Albrigo, L. G. (1980). "Water relations and related leaf characteristics of healthy and blight affected citrus trees." *Journal of the American Society of Horticultural Science,* 105: 431-434.

Szerszen, J. B., Szczyglowski, K. and Bandurski, R. S. (1994). "*iaglu*, a gene from *Zea mays* involved in conjugation of the growth hormone, indole-3-acetic acid (IAA)." *Science,* 265: 1699-1701.

Sztein, A. E., Cohen, J. D., Slovin, J. P. and Cooke, T.J. (1995). "Auxin metabolism in representative land plants." *American Journal of Botany,* 82: 1514–1521.

Sørensen, H. P. & Mortensen, K. K. (2005). "Advanced genetic strategies for recombinant protein expression in *Escherichia coli*." *Journal of Biotechnology,* 115: 113-128.

Sørensen, H. P., Laursen, B. S., Mortensen, K. K. and Sperling-Petersen, H. U. (2002). "Bacterial translation initiation–mechanism and regulation." *Recent Research in Development of Biophysics & Biochemistry,* 2: 243–270.

Takayama, S., Shimosato, H., Shiba, H., Funato, M., Che, F.-S., Watanabe, M., Iwano, M. and Isogai, A. (2001). "Direct ligand-receptor complex interaction controls *Brassica* self-incompatibility." *Nature,* 413: 534-538.

Takayama, S., Shiba, H., Iwano, M., Shimosato, H., Che, F. S., Kai, N., Watanabe, M., Suzuki, G., Hinata, K. and Isogai, A. (2000). "The pollen determinant of self-incompatibility in *Brassica campestris.*" *Proceedings of the National Academy of Sciences USA*, 97: 1920-1925.

Takayanagi, R., Snajdar, R. M., Imada, T., Tamura, M., Pandey, K. N, Misono, K. S. and Inagami, T. (1987). "Purification and characterization of two types of atrial natriuretic factor receptors from bovine adrenal cortex: Guanylate cyclase-linked and cyclase-free receptors." *Biochemical & Biophysical Research Communications,* 144: 244–250.

Takeda, S., Sato, F., Ida, K. and Yamada, Y. (1990). "Characterization of polypeptides that accumulate in cultured *Nicotiana tabacum* cells." *Plant Cell Physiology,* 31: 215–21.

Takei, Y. (2001). "Does the natriuretic peptide system exist throughout the animal and plant kingdom?" *Comparative Biochemistry & Physiology B,* 129: 559–573.

Takei, Y. (2000). "Structural and functional evolution of the natriuretic peptide system in vertebrates." *International Review of Cytology,* 194: 1-66.

Takei, Y., Ueki, M. and Nishizawa, T. (1994). "Eel ventricular natriuretic peptide: cDNA cloning and mRNA expression." *Journal of Molecular Endocrinology*, 13: 339-345.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). "MEGA4: Molecular Evolutionary Analysis (MEGA) software version 4.0." *Molecular Biology & Evolution*, 24: 1596-1599.

Thummler, F. & Verma, D. P. (1987). "Nodulin-100 of soybean is the subunit of sucrose synthase regulated by the availability of free heme in nodules." *Journal of Biological Chemistry*, 262: 14730-14736.

Todd, A., Orengo, C. and Thornton, J. (2001). "Evolution of function in protein superfamilies, from a structural perspective." *Journal of Molecular Biology,* 307(4): 1113-1143.

Topping, J. F. & Lindsey, K. (1997). "Promoter trap markers differentiate structural and positional components of polar development in *Arabidopsis*." *Plant Cell,* 9: 1713–1725.

Tran, L. S., Nakashima, K., Sakuma, Y., Simpson, S. D., Fujita, Y., Maruyama, K., Fujita, M., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2004). "Isolation and functional analysis of arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter." *Plant Cell,* 16: 2481-2498.

Tsurumi, S. & Wada, S. (1986). "Dioxindole-3-acetic acid conjugates formation from indole-3-acetylaspartic acid in *Vicia* seedlings." *Plant Cell Physiology,* 27: 1513–1522.

Tuominen, H., Ostin, A., Sandberg, G. and Sundberg, B. (1994). "A novel metabolic pathway for indole-3-acetic acid in apical shoots of *Populus tremula* (L.) x *Populus tremuloides* (Michx.)." *Plant Physiology,* 106: 1511–1520.

Tyler, R. C., Hassan, S. K., Singh, S., Aceti, D. J., Bingman, C. A., Markley, J. L. and Fox, B. G. (2005). "Auto-induction medium for the production of [U-^{15}N] – and [U-^{13}C, U-^{15}N]-labeled proteins for NMR screening and structure determination." *Protein Expression & Purification,* 40: 268-278.

Urao, T., Yamaguchi-Shinozaki, K., Urao, S. and Shinozaki, K. (1993). "An *Arabidopsis* myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence." *Plant Cell,* 5: 1529-1539.

Valueva, T. A. & Mosolov, V. V. (2004). "Role of inhibitors of proteolytic enzymes in plant defense against phytopathogenic microorganisms." *Biochemistry (Moscow),* 69: 1305-1309.

Van Loon, L. C. (1985). "Pathogenesis-related proteins." *Plant Molecualr Biology,* 4: 111-116.

Van Loon, L. C. & Van Strien, E. A. (1999). "The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins." *Physiological & Molecular Plant Pathology,* 55:85–97.

Van Loon, L. C. & Van Kammen, A. (1970). "Polyacrylamide disc electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var. "Samsun" and "Samsun NN" II. Changes in protein constitution after infection with tobacco mosaic virus." *Virology,* 40: 199-211.

Van Loon, L. C., Rep, M. and Pieterse, C. M. J. (2006). "Significance of inducible defense-related proteins in infected plants." *Annual Review of Phytopathology,* 44: 135–162.

Van Loon, L. C., Pierpoint, W. S., Boller, T. and Conejero, V. (1994). "Recommendations for naming plant pathogenesis related proteins." *Plant Molecular Biology Reporter, 12*: 245-264.

Vesely, D. L. & Giordano, A. T. (1991). "Atrial natriuretic peptide hormonal system in plants." *Biochemical & Biophysical Research Communications,* 179(1): 695–700.

Vesely, M. D., Gower, W. R., Perez-Lamboy, G., Overton, R. M., Graddy, L. and Vesely, D. L. (2001). "Evidence for an atrial natriuretic peptide-like gene in plants." *Experimental Biology and Medicine,* 226: 61-65.

Vesely, D. L., Dietz, J. R., Parks, J. R., Baig, M., McCormick, M. T., Cintron, G. and Schocken, D. D. (1998). "Vessel dilator enhances sodium and water excretion and has beneficial hemodynamic effects in persons with congestive heart failure." *Circulation,* 98: 323–329.

Vesely, D. L., Douglass, M. A., Dietz, J. R., Gower, Jr, W. R., McCormick, M. T., Rodriguez-Paz, G. and Schocken, D. D. (1994). "Three peptides from the atrial natriuretic factor prohormone amino terminus lower blood pressure and produce diuresis, natriuresis, and/or kaliuresis in humans." *Circulation,* 90: 1129-1140.

Vesely, D. L., Gower, Jr, W. R. and Giordano, A. T. (1993). "Atrial natriuretic peptides are present throughout the plant kingdom and enhance solute flow in plants." *American Journal of Physiology*, 265: E465-E477.

Vigers, A. J., Wiedemann, S., Roberts, W. K., Legrand, M., Selitrennikoff, C. and Fritig, B. (1992). "Thaumatin-like pathogenesis-related proteins are antifungal" *Plant Science*, 83:155-161.

Vigers, A. J., Roberts, W. K. and Selitrennikoff, C. P. (1991). "A new family of plant antifungal proteins." *Molecular Plant- Microbe Interactions*, 4: 315-323.

Voesnek, L. A. C. J., Van Der Sman, A. J. M., Harren, F. J. M. and Blom, C. W. P. M. (1992). "An amalgamation between hormone physiology and plant ecology: A review on flooding resistance and ethylene" *Journal of Plant Growth & Regulation*, 11: 171–188.

Voet, D., Voet, J. G. and Pratt, C. W. (1999). "Fundamentals of biochemistry." John Wiley & Sons, Inc., USA.

Vollmar, A. M., Gerbes, A. L., Nemer, M. and Schulz, R. (1993). "Detection of C-type natriuretic peptide (CNP) transcript in the rat heart and immune organs." *Endocrinology*, 132: 1872-1874.

von Heijne, G. (1990). "The signal peptide." *Journal of Membrane Biology*, 115: 195–201.

Waldo, G. S., Standish, B. M., Berendzen, J. and Terwilliger, T. C. (1999). "Rapid protein-folding assay using green fluorescent protein." *Nature Biotechnology*, 17: 691–695.

Walton, D. C. (1980). "Biochemistry and Physiology of Abscisic acid" *Annual Review of Plant Physiology*, 31: 453–489.

Wang, Y. H., Gehring, C., Cahill, D. M. and Irving, H. R. (2007). "Plant natriuretic peptide active site determination and effects on cGMP and cell volume regulation." *Functional Plant Biology*, 34: 645-653.

Wang, D., Amornsiripanitch, N. and Dong, X. (2006). "A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants." *PLoS Pathogens,* 2: e123.

Warneck, H. M., Fulton, D. C., Seitz, H. U. and Fry, S. C. (1998). "Transport, degradation and cell wall-degradation of XXFGol, a growth-regulating nonasaccharide of xyloglucan, in pea stems." *Planta,* 204: 78-85.

Watanabe, M., Ito, A., Takada, Y., Ninomiya, C., Kakizaki, T., Takahata, Y., Hatakeyama, K., Hinata, K., Suzuki, G., Takasaki, T., Satta, Y., Shiba, H., Takayama, S. and Isogai, A. (2000). "Highly divergent sequences of the pollen self-incompatibility (S) gene in class-I S haplotypes of *Brassica campestris* (syn. rapa)." *FEBS Letters,* 473: 139-144.

Watanabe, Y., Barbashov, S. F., Komatsu, S., Hemming, A. M., Miyagi, M., Tsunasawa, A. and Hirano, H. (1994). "A peptide that stimulates phosphorylation of the plant insulin-binding protein: isolation, primary structure and cDNA cloning." *European Journal of Biochemistry,* 224: 167-172.

Wen, J., Lease, K. A. and Walker, J. C. (2004). "DVL, a novel class polypeptides: overexpression alters Arabidopsis development." *Plant Journal,* 37: 668-677.

Werner, T., Fessele, S., Maier, H. and Nelson, P. J. (2003). "Computer modeling of promoter organisation as a tool to study transcriptional coregulation." *FASEB Journal,* 17: 1228-1237.

Westbrook, J., Feng, Z., Chen, L., Yang, H. and Berman, H. M. (2003). "The Protein Data Bank and structural genomics." *Nucleic Acids Research,* 31: 489–491.

Woloshuk, C. P., Meulenhoff, J. S., Sela-Buurlage, M., van den Elzen, P. J. M. and Cornelissen, B. J. C. (1991). "Pathogen-induced proteins with inhibitory activity toward *Phytophthora infestans.*" *The Plant Cell,* 3: 619-628.

Wu, F. & Wu, Q. (2003). "Corin-mediated processing of pro-atrial natriuretic peptide in human small cell lung cancer cells." *Cancer Research,* 63: 8318–8322.

Wu, J. & Filutowicz, M. (1999). "Hexahistidine (His6)-tag dependent protein dimerization: a cautionary tale." *Acta Biochimica Polonica,* 46: 591–599.

Wu, C., Wu, F., Pan, J., Morser, J. and Wu, Q. (2003). "Furin-mediated processing of pro-C-type natriuretic peptide." *Journal of Biological Chemistry,* 278(28): 25847–25852.

Wu, C-T., Leubner-Metzger, G., Meins, F. and Bradford, K. J. (2001). "Class I β-1,3-glucanase and chitinase are expressed in the micropylar endosperm of tomato seeds prior to radicle emergence." *Plant Physiology,* 126: 1299–1313.

Xiong, L., Schumaker, K. S. and Zhu, J-K. (2002). "Cell signaling during cold, drought, and salt stress." *The Plant Cell,* 14: S165–S183.

Yalpani, N. & Raskin, I. (1993). "Salicylic acid: A systematic signal in induced plant decrease resistance." *Trends in Microbiology,* 1: 88-92.

Yalpani, N., Silverman, P., Wilson, T. M. A., Kleier, D. A. and Raskin, I. (1991), "Salicylic acid is a systemic signal and an inducer of pathogenesis-related proteins in virus-infected tobacco." *Plant Cell,* 3: 809–818.

Yamakawa, S., Matsubayashi, Y., Sakagami, Y., Kamada, H. and Satoh, S. (1999). "Promotive effects of the peptidyl plant growth factor, phytosulfokine-α, on the growth and chlorophyll content of *Arabidopsis* seedlings under high night-time temperature conditions." *Bioscience, Biotechnology & Biochemistry,* 63: 2240-2243.

Yamakawa, S., Sakurai, C., Matsubayashi, Y., Sakagami, Y., Kamada, H. and Satoh, S. (1998). "The promotive effects of a peptidyl plant growth factor, phytosulfokine, on the formation of adventitious roots and expression of a gene for a root-specific cystatin in cucumber hypocotyls." *Journal of Plant Research,* 111: 453-458.

Yamazaki, T., Takaoka, M., Katoh, E., Hanada, K., Sakita, M., Sakata, K., Nishiuchi, Y. and Hirano, H. (2003). "A possible physiological function and the tertiary structure of a 4-kDa peptide in legumes." *European Journal of Biochemistry,* 270:1269-1276.

Yan, Z., Reddy, M. S., Yyu, C.-M., McInroy, J. A., Wilson, M. and Kloepper, J. W. (2002). "Induced systemic protection against tomato late blight by plant growth-promoting rhizobacteria." *Phytopathology,* 92: 1329-1333.

Yang, G., Shen, S., Kobayashi, T., Matsubayashi, Y., Sakagami, Y. and Kamada, H. (1999a). "Stimulatory effects of a novel peptidyl plant growth factor, phytosulfokine-α, on adventitious bud formation in *Antirrhinum majus*." *Plant Biotechnology,* 16: 231-234.

Yang, H., Matsubayashi, Y., Nakamura, K. and Sakagami, Y. (1999b). "*Oryza sativa* PSK encodes a precursor of phytosulfokine, a sulfated peptide growth factor found in plants." *Proceedings of the National Academy of Sciences USA*, 96: 13560-13565.

Yang, W., Katinakis, P., Hendriks, P., Smolders, A., de Vries, F., Spee, J., van Kammen, A., Bisseling, T. and Franssen, H. (1993). "Characterization of *GmENOD40*, a gene showing novel patterns of cell-specific expression during soybean nodule development." *Plant Journal*, 3(4): 573–585.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors." *Gene,* 33: 103-119.

Yu, D., Chen, C. and Chen, Z. (2001). "Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression." *Plant Cell*, 13: 1527-1540.

Zeevart, J. A. D. and Creelman, R. A. (1988). "Metabolism and physiology of abscisic acid" *Annual Review of Plant Physiology and Plant Molecular Biology,* 39: 439 -473.

Zeevart, J. A. D., Rock, C. D., Fantauzzo, F., Heath, T. G. and Gage, D. A. (1991). "Metabolism of abscisic acid and its physiological implications" In Abscisic Acid:

Physiology and Biochemistry, W. J. Davies and H. G. Jones, eds. BIOS Scientific, Oxford, UK, pp 39–52.

Zenoni, S., Reale L., Tornielli, G.B., Lanfaloni, L., Porceddu, A., Ferrarini, A., Moretti, C., Zamboni, A., Speghini, A., Ferranti, F. and Pezzotti, M. (2004). "Downregulation of the Petunia hybrida a-expansin gene PhEXP1 reduces the amount of crystalline cellulose in cell walls and leads to phenotypic changes in petal limbs." *Plant Cell*, 16: 295–308.

Zhu, J. K. (2002). "Salt and drought stress signal transduction in plants." *Annual Review of Plant Biology*, 53: 247-273.

Zomorrodipour, A., Yakhchali, B., Khodabandeh, M., Deezagi, A., Hosseini Mazinani, S. M., Valian Borujeni, S., Ahmadzadeh Raji, M., Rahimi, M., Ahmadi Danesh, H. and Sanati, M. H. (2004). "The over-expression of biologically active human growth hormone in a T5-based system in *Escherichia coli*, studying temperature effect." *Journal of science, Islamic Republic of Iran*, 15(1): 1016-1104.

Websites

http://au.expasy.org/prosite/

http://bioinf.cs.ucl.ac.uk/psipred/

http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

http://blast.ncbi.nlm.nih.gov/Blast.cgi

http://www.cbs.dtu.dk/services/promoter

http://www.cbs.dtu.dk/services/TargetP/

http://compbio.dfci.harvard.edu/tgi/

http://www.dna.affrc.go.jp/PLACE/signalscan.html

http://www.ebi.ac.uk/Tools/clustalw2/index.html

http://www.fruitfly.org/seq_tools/promoter.html

http://www.genomatix.de/cgi-bin/matinspector_prof/mat_fam.pl?

http://www.ncbi.nlm.nih.gov/Entrez

http://www.ncbi.nlm.nih.gov/projects/sviewer

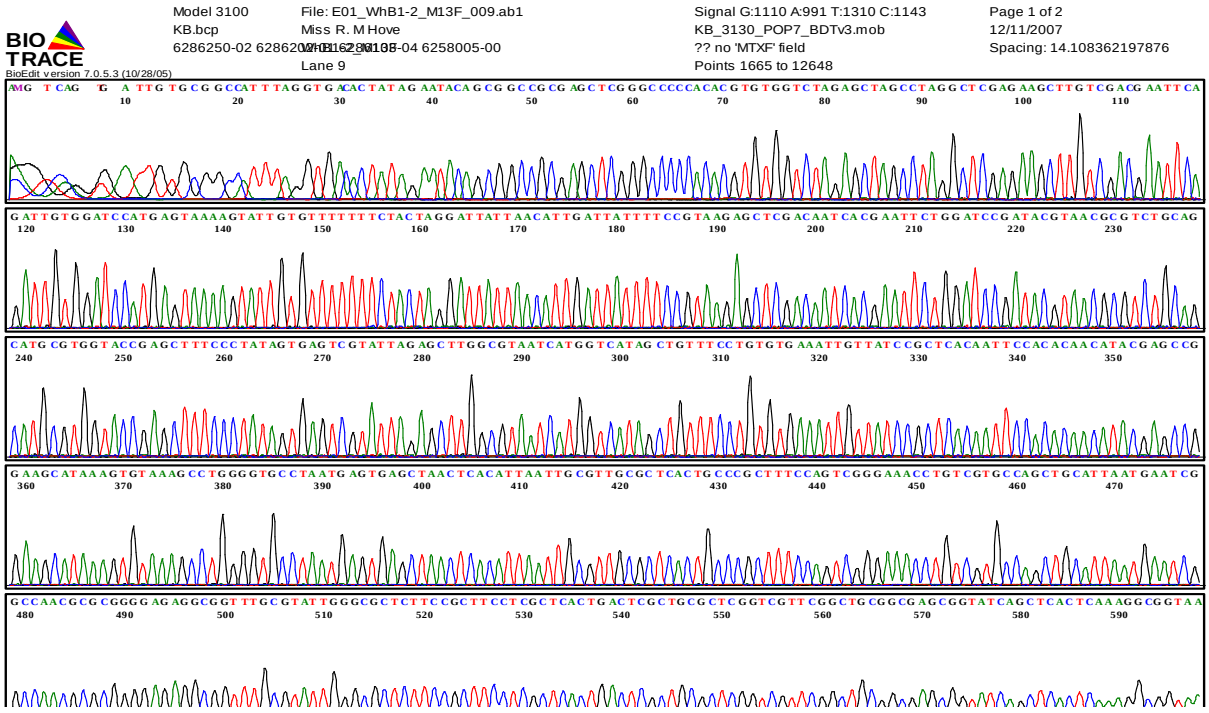
http://www1.qiagen.com/literature/pDrive/pcr_cloning21.pdf

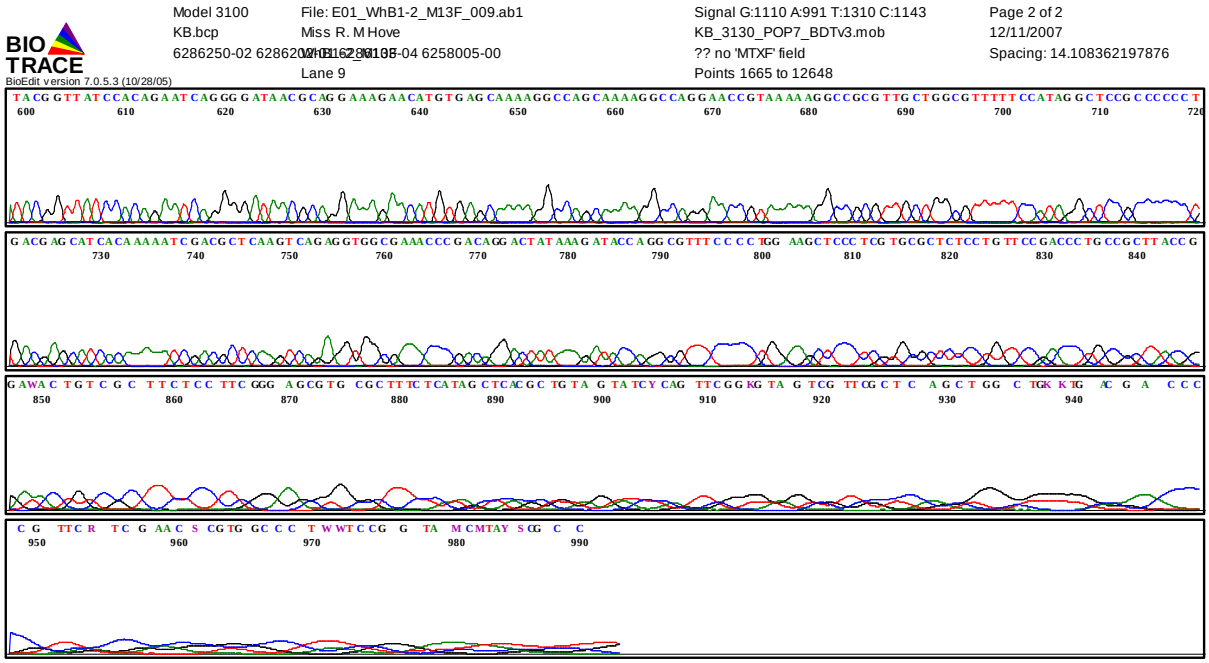
http://swissmodel.expasy.org/

APPENDICES

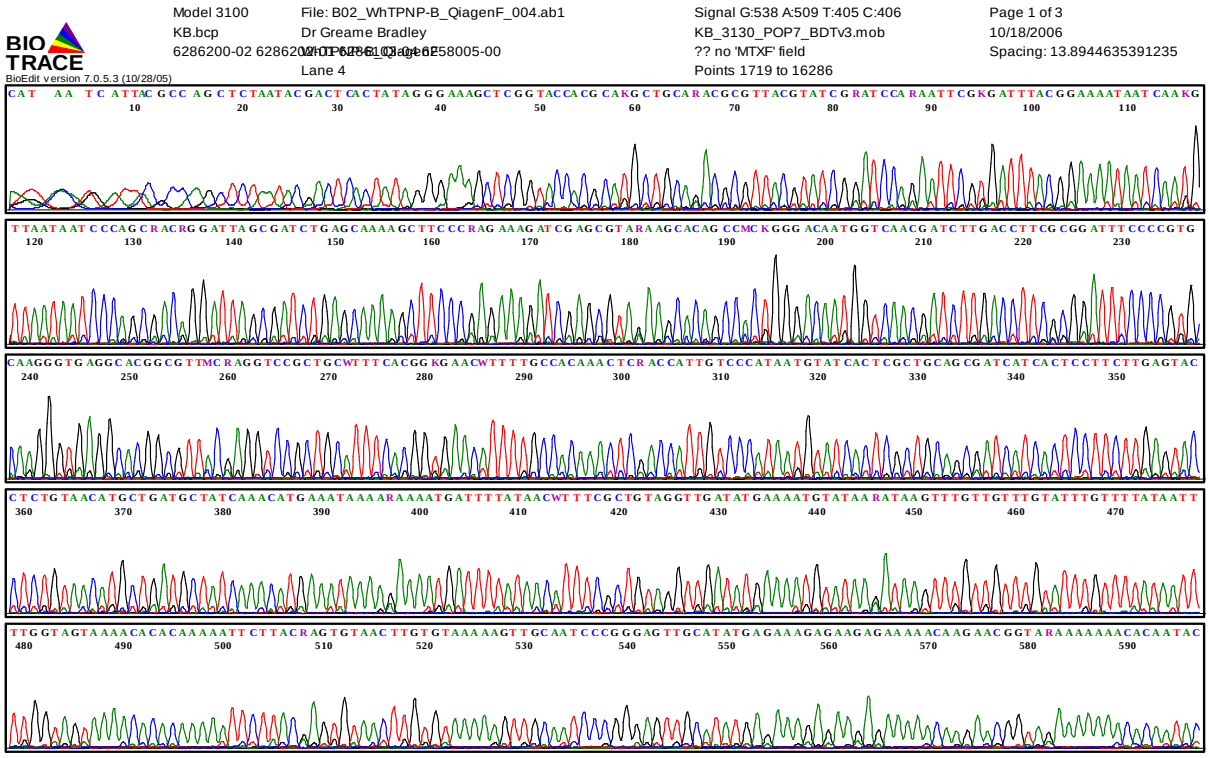
APPENDIX 1: Sequence Chromatograms

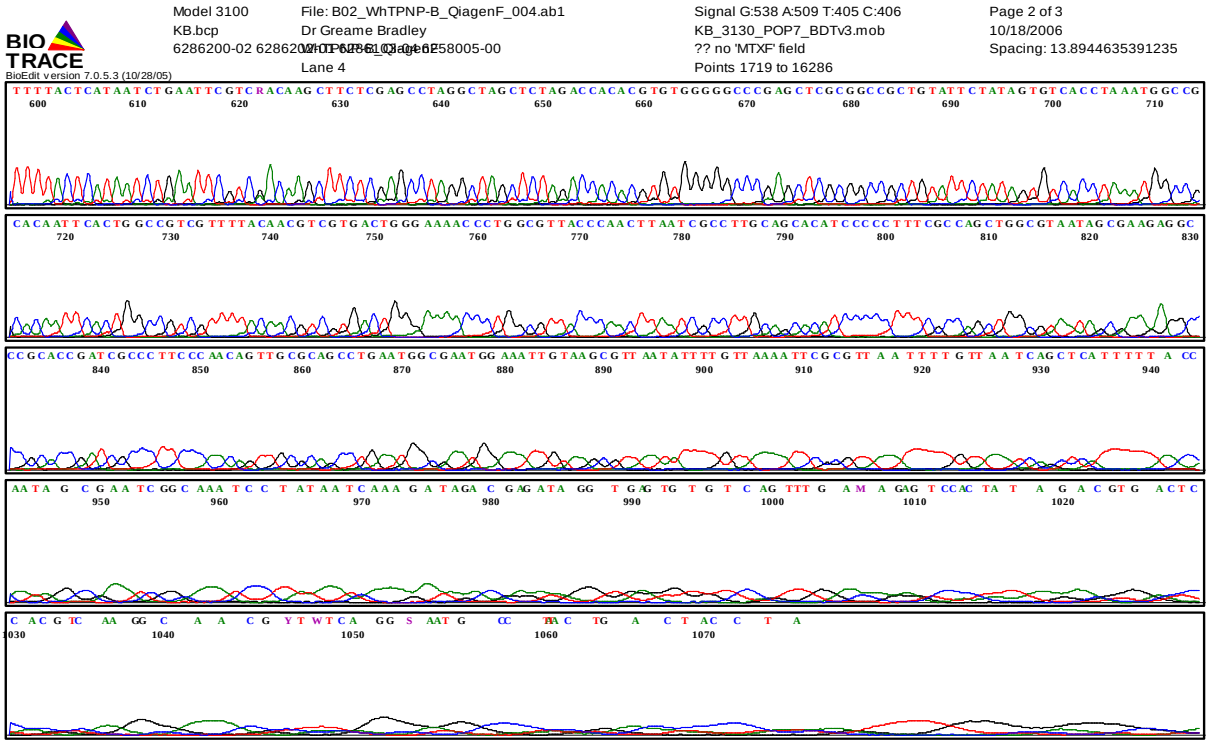
Wheat (Betta) – M13 Forward





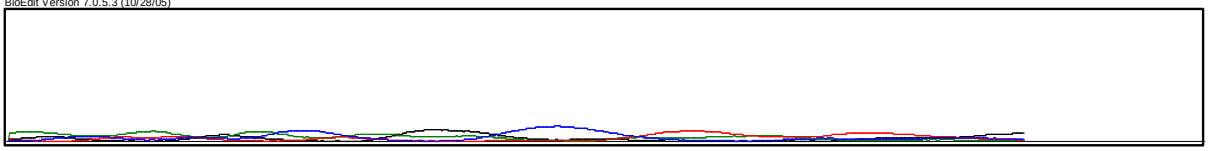
Wheat (Tugella) – Qiagen Forward



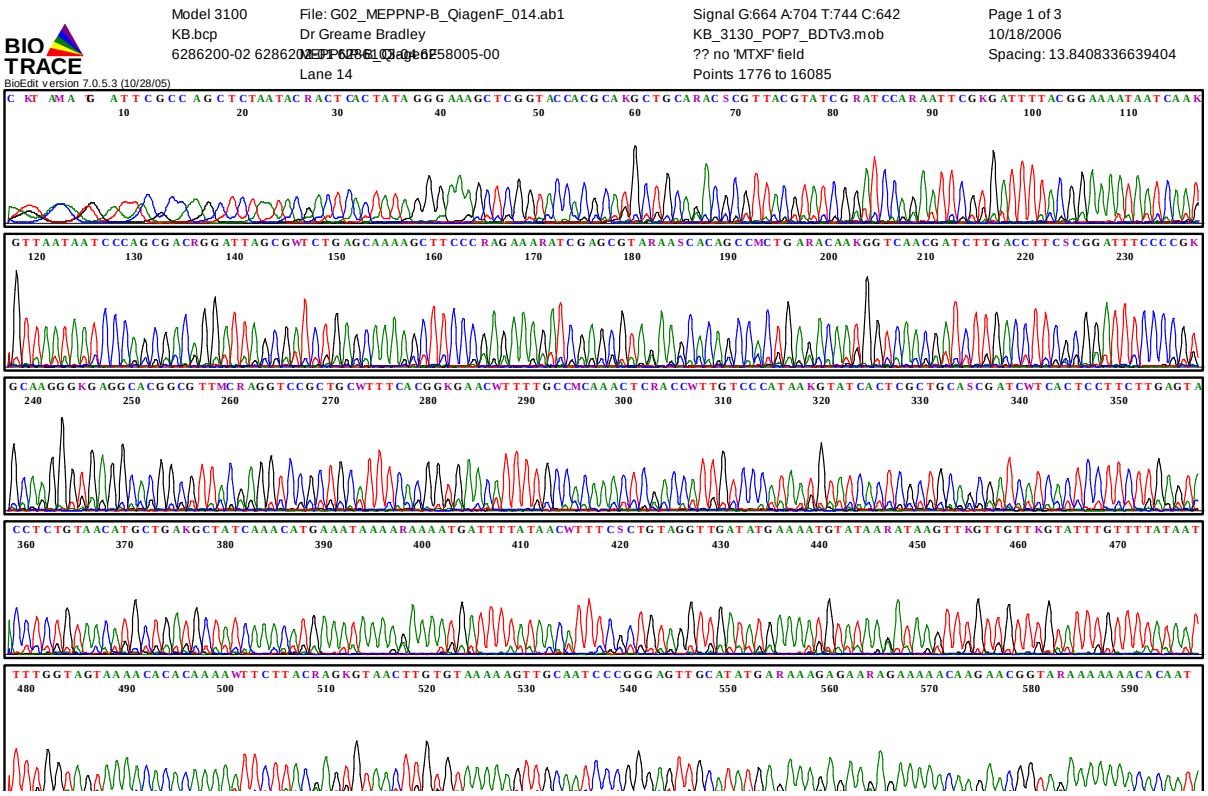


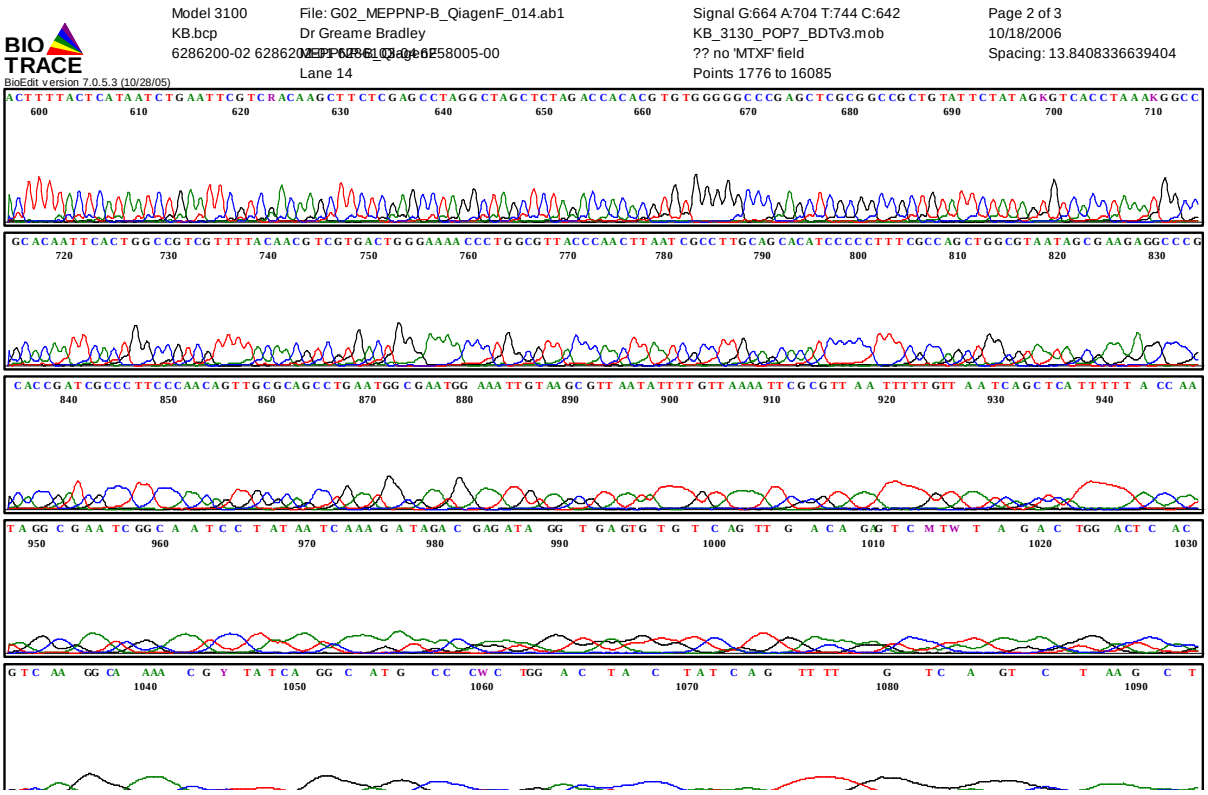
BIO TRACE
BioEdit version 7.0.5.3 (10/28/05)

Model 3100	File: B02_WhTPNP-B_QiagenF_004.ab1	Signal G:538 A:509 T:405 C:406	Page 3 of 3
KB.bcp	Dr Greame Bradley	KB_3130_POP7_BDTv3.mob	10/18/2006
6286200-02 6286200-02_WhTPNP-B_QiagenF 258005-00		?? no 'MTXF' field	Spacing: 13.8944635391235
	Lane 4	Points 1719 to 16286	



Maize (early pearl) – Qiagen Forward





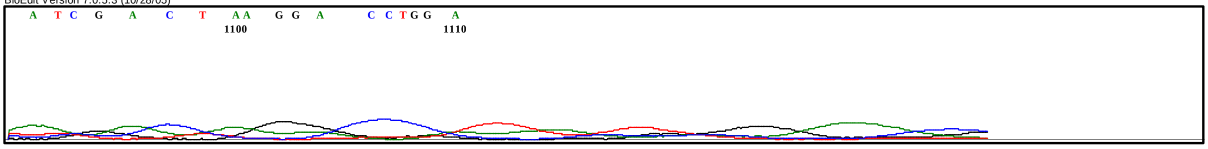
BIO TRACE
BioEdit version 7.0.5.3 (10/28/05)

Model 3100
KB.bcp
6286200-02 62862MEPPNP-B_QiagenF58005-00

File: G02_MEPPNP-B_QiagenF_014.ab1
Dr Greame Bradley
Lane 14

Signal G:664 A:704 T:744 C:642
KB_3130_POP7_BDTv3.mob
?? no 'MTXF' field
Points 1776 to 16085

Page 3 of 3
10/18/2006
Spacing: 13.8408336639404



APPENDIX 2: Transcription Binding Factors

Transcription Factor	Consensus Sequence	Description	Species	Reference
ABRELATERD1	ACGTG	ABRE-like sequence (from -199 to -195) required for etiolation-induced expression of erd1 (early responsive to dehydration) in *Arabidopsis*.	*Arabidopsis thaliana*	Nakashima *et al.*, 2006
ABREOSRAB21	ACGTSSSC	ABA responsive element (ABRE) of wheat Em and rice (O.s.) rab21 genes.	rice (*Oryza sativa*); wheat (*Triticum aestivum*)	Busk & Pages, 1998
ACGTATERD1	ACGT	ACGT sequence (from -155 to -152) required for etiolation-induced expression of erd1 (early responsive to dehydration) in *Arabidopsis*.	*Arabidopsis thaliana*	Simpson *et al.*, 2003
CAATBOX1	CAAT	CAAT promoter consensus sequence found in legA gene of pea.	Pea (*Pisum sativum*)	Shirsat *et al.*, 1989
ERELEE4	AWTTCAAA	"ERE (ethylene responsive element)" of tomato (L.e.) E4 and carnation GST1 genes; GST1 is related to senescence; Found in the 5'-LTR region of TLC1.1 retrotransposon family in Lycopersicon chilense; ERE motifs mediate ethylene-induced activation of the U3 promoter region.	tomato (*Lycopersicon esculentum*); carnation (*Dianthus aryophillus*); *Lycopersicon chilense*.	Rawat *et al.*, 2005
GCCCORE	GCCGCC	Core of GCC-box found in many pathogen-responsive genes such as PDF1.2, Thi2.1, and PR4; Has been shown to function as ethylene-responsive element; Appears to play important roles in regulating jasmonate-responsive gene expression; Tomato Pti4 (ERF) regulates defence-related gene expression via GCC box and non-GCC box cis elements (Myb1 (GTTAGTT) and G-box (CACGTG).	*Arabidopsis thaliana*; *Lycopersicon esculentum* (tomato)	Koyama *et al.*, 2003

INRNTPSADB	YTCANTYY	Inr (initiator) elements found in the tobacco psaDb gene promoter without TATA boxes; Light-responsive transcription of psaDb depends on Inr, but not TATA box.	Tobacco (*Nicotiana tabacum*)	Nakamura *et al.*, 2002
LTRECOREATCOR15	CCGAC	Core of low temperature responsive element (LTRE) of cor15a gene in Arabidopsis; A portion of repeat-C (C-repeat), TGGCCGAC, which is repeated twice in cor15a promoter (Baker *et al.*, 1994); ABA responsiveness; Involved in cold induction of BN115 gene from winter *Brassica napus*; Light signaling mediated by phytochrome is necessary for cold- or drought-induced gene expression through the C/DRE in *Arabidopsis.*	*Arabidopsis thaliana*; *Brassica napus*	Baker *et al.*, 1994
MYB1AT	WAACCA	MYB recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in *Arabidopsis.*	*Arabidopsis thaliana*	Abe *et al.*, 2003
MYB2AT	TAACTG	Binding site for ATMYB2, an *Arabidopsis* MYB homolog; ATMYB2 binds oligonucleotides that contained a consensus MYB recognition sequence (TAACTG), such as is in the SV40 enhancer and the maize bronze-1 promoter (Urao *et al.*, 1993); ATMYB2 is involved in regulation of genes that are responsive to water stress in *Arabidopsis*.	*Arabidopsis thaliana*	Urao *et al.*, 1993
MYB2CONSENSUSAT	YAACKG	MYB recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in *Arabidopsis.*	*Arabidopsis thaliana*	Abe et al., 2003

MYBCORE	CNGTTR	Binding site for all animal MYB and at least two plant MYB proteins ATMYB1 and ATMYB2, both isolated from *Arabidopsis*; ATMYB2 is involved in regulation of genes that are responsive to water stress in Arabidopsis; A petunia MYB protein (MYB.Ph3) is involved in regulation of flavonoid biosynthesis.	*Arabidopsis thaliana;* animal; petunia (*Petunia hybrida*)	Solano *et al.*, 1995
MYCATERD1	CATGTG	MYC recognition sequence (from -466 to -461) necessary for expression of erd1 (early responsive to dehydration) in dehydrated Arabidopsis; NAC protein bound specifically to the CATGTG motif (Tran *et* al., 2004); NAC protein bound specifically to the CATGTG motif (Tran *et al.*, 2004).	*Arabidopsis thaliana*	Tran *et al.*, 2004
MYCATRD22	CACATG	Binding site for MYC (rd22BP1) in *Arabidopsis* dehydration-resposive gene, rd22; MYC binding site in rd22 gene of *Arabidopsis thaliana*; ABA-induction; Located at ca. -200 of rd22 gene; Also MYB at ca. -141 of rd22 gene.	*Arabidopsis thaliana*	Abe *et al.*, 1997
MYCCONSENSUS AT	CANNTG	MYC recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in Arabidopsis; Binding site of ATMYC2 (previously known as rd22BP1); MYC recognition sequence in CBF3 promoter; Binding site of ICE1 (inducer of CBF expression 1) that regulates the transcription of CBF/DREB1 genes in the cold in *Arabidopsis*.	*Arabidopsis thaliana*	Abe *et al.*, 2003
POLASIG3	AATAAT	"Plant polyA signal"; Consensus sequence for plant polyadenylation signal.	Maize (*Zea mays*)	Heidecker & Messing, 1986

PYRIMIDINEBOX HVEPB1	TTTTTTCC	“Pyrimidine box" found in the barley (H.v.) EPB-1 (cysteine proteinase) gene promoter; Located between -120 to -113; Required for GA induction.	Barley (*Hordeum vulgare*)	Cercos *et al.*, 1999
TAAAGSTKST1	TAAAG	TAAAG motif found in promoter of Solanum tuberosum KST1 gene; Target site for trans-acting StDof1 protein controlling guard cell-specific gene expression; KST1 gene encodes a K+ influx channel of guard cells.	Potato (*Solanum tuberosum*)	Plesch *et al.*, 2001
TATABOX1	CTATAAATAC	"TATA box"; TATA box found in the 5'upstream region of rice alpha-amylase; TATA box found in beta-phaseolin promoter (Grace *et al.*, 2004); sequence and spacing of TATA box elements are critical for accurate initiation (Grace *et al.*, 2004).	rice (*Oryza sativa*); bean (*Phaseolus vulgaris*)	Grace *et al.*, 2004
TATABOX2	TATAAAT	"TATA box"; TATA box found in the 5'upstream region of pea legA gene; sporamin A of sweet potato; TATA box found in beta-phaseolin promoter (Grace *et al.*, 2004); sequence and spacing of TATA box elements are critical for accurate initiation (Grace *et al.*, 2004).	pea (*Pisum sativum*); tobacco (*Nicotiana tabacum*); bean (*Phaseolus vulgaris*)	Shirsat *et al.*, 1989 Grace *et al.*, 2004
TATABOX4	TATATAA	"TATA box"; TATA box found in the 5'upstream region of sweet potato sporamin A gene; TATA box found in beta-phaseolin promoter (Grace *et al.*, 2004); sequence and spacing of TATA box elements are critical for accurate initiation (Grace *et al.*, 2004).	sweet potato (*Ipomoea batatas*); bean (*Phaseolus vulgaris*)	Grace *et al.*, 2004

WBOXATNPR1	TTGAC	W-box found in promoter of *Arabidopsis thaliana* NPR1 gene; Located between +70 and +79 in tandem; They were recognized specifically by salicylic acid (SA)-induced WRKY DNA binding proteins; A cluster of WRKY binding sites act as negative regulatory elements for the inducible expression of AtWRKY18 (Chen and Chen, 2002).	*Arabidopsis thaliana*	Yu *et al.*, 2001

The symbols used in addition to A, G, C, or T are:

N: A, C, G or T;

R: A or G;

W: A or T;

Y: C or T

S: C or G