Recombinant Expression and Functional Characterisation of a Putative Clathrin Assembly Protein from *Arabidopsis thaliana*

P Chatukuta

(23176466)

Thesis submitted in fulfillment of the requirements for the degree of **Doctor of Philosophy** in Biology at the Mafikeng Campus of the North-West University

Promoter : Prof. O Ruzvidzo

June 2015

Declaration

I, Patience Chatukuta, declare that the thesis entitled "**Recombinant Expression and Functional Characterisation of a Putative Clathrin Assembly Protein from *Arabidopsis thaliana***" is my work and has not been submitted for any degree or examination at any other university or institution and that all sources of my information here have been acknowledged as indicated in the text and/or list of references.

Student: Patience Chatukuta

Signature: ..... **Date:** .....

Promoter: Prof. O Ruzvidzo

Signature: ..... **Date:** .....

Dedication

This thesis is dedicated to the phenomenal women scientists from the developing countries who have overcome the seemingly insurmountable odds to contribute towards scientific research and development. Continue to enlighten!!!

Acknowledgements

This thesis would not have been possible were it not for the assistance and support of various people, of whom there are too many to mention. However, the following few deserve a special mention:

My supervisor and mentor, Professor Oziniel Ruzvidzo, had a vision and this thesis is the manifestation of that vision. His passion for biotechnology and science in general is infectious and I caught the bug. I appreciate his vision, hardwork, tenacity and guidance. I will always maintain that he is the best supervisor in the world.

Professor Chris Gehring laid the foundation for the study of adenylate cyclases in Arabidopsis using modern biotechnology techniques. We all have to start somewhere and I am thankful that his work provided me with a basis to add a few bricks to this building of plant nucleotide research.

I am grateful to Doctor Lusisizwe Kwezi for teaching me many of the molecular biology techniques that enabled me to meet the objectives of this research study.

The members of staff of the North-West University's Department of Biological Science, in particular, Dr. Dave Kawadza, Ms. Bridget Tshegofatso Dikobe, Ms. Rika Huyser and Ms. Tebogo Mabunda, were extremely forthcoming whenever I approached them for assistance.

I value the support of the NWU Plant Biotechnology Research Group. Their sense of humour made the long hours in the laboratory more bearable.

My friend, Grace Wanjiku Migwi Chivaura, played a pivotal role in my life by introducing me to Prof. Ruzvidzo five years ago and thus fostering a very fruitful relationship.

This research study was funded by the North-West University (NWU) – Mafikeng Campus, the South African National Research Foundation (NRF), the Organization for Women in Science for the Developing World (OWSD) and the Swedish International Development Cooperation Agency (SIDA). Their financial assistance enabled the progress of this research study to culminate and significantly contributed towards the development of skills in the field of biotechnology, which is recognized as a scarce skills and priority area for the Southern African region.

I also express my gratitude to the Mafikeng Assembly, led by Pastor M. Phaladi, who became my family away from home, and provided moral support and encouragement. I am especially grateful to MaPhaladi, Mrs D Phaladi, who took it upon herself to be a mother, friend and confidante. I appreciate her greatly.

My families, the Chatukutas and the Kwangwas, have been a pillar of support. They have allowed me to take my own road in life and cheered me on. I thank my husband, Edi Kwangwa, for putting up with some absences from the family home and being my most prominent supporter. I love you all.

May God bless them all abundantly!!!!!

Above all, I am grateful to the Almighty God, the God of mercy, the God of second chances. All glory is due to Him for His enabling grace.

List of Abbreviations

AC: Adenylate cyclase

ANOVA: Analysis of variance

ANTH: AP180 N-terminal homology

AP: Assembly protein

AtBRI1: *Arabidopsis thaliana* brassinosteroid receptor 1

AtCAP: *Arabidopsis thaliana* clathrin assembly protein

AtCAP-AC: *Arabidopsis thaliana* clathrin assembly protein adenylate cyclase domain fragment

AtGC1: *Arabidopsis thaliana* guanylate cyclase 1

AtNOGC1: *Arabidopsis thaliana* nitric oxide-binding guanylate cyclase

ATP: Adenosine 5′-triphosphate

AtPepR1: *Arabidopsis thaliana* pattern recognition receptor

AtPPR: *Arabidopsis thaliana* pentatricopeptide protein

AVT: AtGenExpressVisualisation Tool

BioGRID: Biological General Repository for Interaction Datasets

BLAST: Basic Local Alignment Search Tool

BLASTP: Protein-protein Basic Local Alignment Search Tool

bp: Base pairs

CALM: Clathrin assembly lymphoid myeloid leukemia

cAMP: Cyclic adenosine 3′,5′-monophosphate

cAMP-CAP: cAMP-catabolite activator protein complex

CAP: Clathrin assembly protein

CCV: Clathrin-coated vesicle

c-diGMP: Cyclic bis-(3′,5′) di-guanosine monophosphate

cDNA: Copy DNA

CDS: Coding DNA sequence

CGAP: Catabolite gene-activator protein

cGMP: Cyclic guanosine 3′,5′-monophosphate

cNMP: Cyclic nucleotide monophosphate

CREB: cAMP-response element-binding

C-terminal: Carboxy terminal

cv: Column volume

***cya*:** Adenylate cyclase gene

CYTH: *Cya*B, thiamine triphosphatase

DNA: Deoxyribonucleic acid

E: Expectation

EDTA: Ethylene di-amine tetra-acetic acid

ENTH: Epsin N-terminal homology

FAO: Food and Agriculture Organisation

Fve: Fungal immunomodulatory protein fve

GAT: GGA and Tom1

GC: Guanylate cyclase

GEO: Gene Expression Omnibus

GGA: Golgi-localising, gamma-adaptin ear domain homology, ARF-binding protein

G-protein: Guanine nucleotide-binding protein

GTP: Guanosine triphosphate

GTPase: Guanosine triphosphatase

HpAC1: *Hippeastrum hybridum* adenylate cyclase 1

IBMX: 3-Isobutyl-1-methylxanthine

IPTG: Isopropyl β-D-1-thiogalactopyranoside

kD: KiloDalton

KEGG: Kyoto Encyclopedia of Genes and Genomes

LB: Luria-Bertani

MOPS: 3-(N-morpholino) propanesulfonic acid

mRNA: Messenger ribonucleic acid

MS: Murashige and Skoog

NASC: Nottingham Arabidopsis Stock Centre

NbAC: *Nicotiana benthamiana* adenylate cyclase

NCBI: National Centre for Biotechnology Information

Ni-NTA: Nickel-nitrilotriacetic acid

N-terminus: Amino-terminus

OmpT: Outer membrane protease

ORF: Open reading frame

P: Probability

PCR: Polymerase Chain Reaction

PHYRE2: Protein Homology/AnalogY Recognition Engine 2

PIR: Protein Information Resource

PMSF: Phenylmethylsulfonyl fluoride

PPDB: Plant Proteome Database

PSiP: Pollen signaling protein

PSKR1: Phytosulfokine receptor 1

RefSeq: NCBI Reference Sequence Database

RNA: Ribonucleic acid

RNAse: Ribonuclease

ROS: Reactive oxygen species

RT-PCR: Reverse Transcriptase - Polymerase Chain Reaction

sAC: Soluble adenylate cyclase

SDS: Sodium dodecyl sulphate

SDS-PAGE: Sodium dodecyl sulphate - polyacrylamide gel electrophoresis

SNK: Student Newman-Kuehls

SOC: Super optimal broth with catabolite repression

SORLIP: Sequences Over-represented in Light-induced Promoter

SSDB: Sequence Similarity Database

STRING: Search Tool for Recurring Instances of Neighbouring Genes

SUBA3: The Sub-cellular Proteomic Database

TAIR: The Arabidopsis Information Resource

TBE: Tris/borate/EDTA

TBS: Tris-buffered saline

TFBS: Transcription factor binding site

TIGR: The Institute for Genomic Research

TIR: Toll/interleukin-1 receptor

TIR-NBS-LRR: Toll interleukin-like receptor nucleotide binding site leucine-rich repeat

tmAC: Trans-membrane adenylate cyclase

tRNA: Transfer ribonucleic acid

tRNA_SAD: tRNA synthetase second additional domain

TSS: Transcription start site

UN: United Nations

UniProt: Universal Protein Resource

UniProtKB: UniProt Knowledge Base

UV: Ultraviolet

VHS: VPS-27, Hrs and STAM

Key Terms

Abiotic Stress: The negative impact of non-living factors on other living organisms in a specific environment.

Adenylate Cyclase: An enzyme that synthesizes cyclic adenosine monophosphate (cAMP) from adenosine triphosphate (ATP).

Annotation: A determination of the locations of genes, their coding regions in a genome and the functions of such genes.

Bioinformatics: The collection, classification, storage, and analysis of biological information using computers.

Biotic Stress: A stress aspect that occurs to plants as a result of damage done to them by other living organisms.

Clathrin Assembly Protein: A protein that stimulates the assembly of clathrin lattices onto cellular membranes.

Climate Change: The long-term shift in weather patterns in a specific region or globally.

Cloning: A process of producing similar copies of genetically identical materials and/or individuals.

Co-expression: The simultaneous expression of two or more genes.

Complementation: A genetic cross used to identify if two mutations are located within the same or different gene.

Cyclic Adenosine 3′,5′-monophosphate (cAMP): A second messenger cyclic nucleotide formed from adenosine triphosphate by the action of the enzyme adenylate cyclase that participates in signal transduction.

Domain: A conserved part of a given protein sequence and structure that can evolves, functions, and exists independently of the rest of the protein chain.

Endocytosis: A process by which cells absorb molecules (such as proteins) by engulfing.

Endogenous: Originating from within an organism's tissues or cells.

Enzyme Immunoassay: A test that combines antibody binding with enzymatic detection to quantify biological molecules of interest.

Epsin N-terminal Homology (ENTH): A structural domain found in proteins that is involved in endocytosis and the cytoskeletal machinery.

Expression Construct: A plasmid that carries and transmits specific genes for expression within target cells.

Expression Profile: A description of the (relative) expression levels of a gene across a set of experimental conditions.

Expressional Analysis: The determination of the pattern of genes expressed at the level of genetic transcription, under specific circumstances or in a specific cell.

Heterologous Expression: The expression of a gene or part of a gene in a host organism, which naturally does not have this gene or gene fragment.

***In Vitro*:** Occurring or made to occur in a laboratory vessel or other controlled experimental environments rather than within a living organism or natural setting.

***In Vivo*:** Occurring or made to occur within a living organism or natural setting.

MacConkey Lactose Agar: A selective and differential medium designed to isolate and differentiate enterics based on their ability to ferment lactose.

Microarray: A set of DNA/RNA sequences representing the entire set of genes of an organism, arranged in a grid pattern for use in genetic testing.

Motif: A sequence pattern of nucleotides in a DNA molecule or amino acids in a protein molecule.

Polymerase Chain Reaction (PCR): A technique of amplifying a single copy or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of that particular DNA sequence.

Primer: A strand of short nucleic acid sequences, generally about 10 base pairs, that serves as a starting point for DNA/RNA synthesis.

Promoter: A region of DNA that initiates transcription of a particular gene.

Re-folding: A process that restores the biological activity of an unfolded or misfolded protein.

Sodium Dodecyl Sulphate - Polyacrylamide Gel Electrophoresis (SDS-PAGE): The separation of proteins according to their molecular weights, and based on their differential rates of migration through a sieving gel matrix under the influence of an applied electrical field.

Second Messenger: Molecules that relay signals received at receptors on the cell surface to target molecules within the cytosol and/or nucleus.

Signal Transduction: The transmission of molecular signals from a cell's exterior to its interior.

Transformation: The genetic alteration of a cell resulting from the direct uptake and incorporation of exogenous genetic material.

Table of Contents

List of Figures

List of Tables

Overall Research Summary

The need to develop agricultural crops with persistent resistance to the vagaries of climate change has led plant biotechnologists to strategically focus on those plant molecules involved in the maintenance and sustenance of homeostasis. One such plant molecule that is typically involved in the various signal transduction and cellular communication systems is the cyclic adenosine 3′,5′-monophosphate (cAMP) which naturally, is generated by the enzyme, adenylate cyclase (AC). Even though ACs have previously been experimentally proven to be centrally involved in numerous stress response systems in animals, prokaryotes and lower eukaryotes, their existence and/or functional properties in higher plants have until recently, been a very serious matter of debate and elusiveness. To date, only four ACs have since been confirmed in higher plants, and specifically in the *Arabidopsis thaliana, Zea mays, Nicotiana benthamiana* and *Hippeastrum hybridum* plants. Since it is inconceivable that a single AC per plant can account for all the cAMP-dependent processes in plants, we set out to enzymatically and functionally characterize a second probable AC candidate from *A. thaliana* in the form of a putative clathrin assembly protein (AtCAP: At1g68110), with a view of elucidating its exact physiological and biological roles in higher plants. From our findings, a preliminary bioinformatic analysis of this protein showed that this putative candidate is actually a multi-domain, multi-functional protein with a possible role in AC-dependent stress response and adaptation mechanisms. We then cloned and recombinantly expressed its AC-domain-containing fragment of AtCAP (AtCAP-AC) in the chemically competent *E. cloni* BL21 (DE3) pLysS DUOs cells and unequivocally demonstrated its ability to induce the generation of endogenous cAMP in this prokaryotic expression system. More so, we also undoubtedly

demonstrated a complementation system of the mutant non-lactose fermenting *cya*A SP850 *Escherichia coli* strain by this recombinant AtCAP-AC protein, to eventually ferment lactose and as a result of this protein's ability to generate the most required cAMP molecule for this process. In addition, we also undertook the purification system of this truncated AtCAP-AC protein followed by the functional characterization of its *in vitro* AC activities, which in turn revealed that this recombinant is indeed a *bona fide* soluble adenylate cyclase (sAC), whose physiological and biochemical roles may be mediated by cAMP and possibly, via a calmodulin-dependent signalling system. Lastly and in order to gain more insights into the possible physiological roles of this AtCAP-AC protein in higher plants, we then bioinformatically analysed its expressional profiles using the various computational and web-based bioinformatic tools, and found out that the protein is centrally involved in responses to biotic stress factors, whose systems are both cAMP- and SORLIP1AT core motif-dependent.

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 INTRODUCTION

1.1.1 Background

Plants play a very important and central role in the earth's ecosystem because they produce food that is later on made available to consumers such as humans and animals. As such, the farming of crop plants for food and feed purposes makes agriculture a vital system to the sustainable existence of both humankind and the animals that solely depend on it. However, the productivity of crop plants is somewhat becoming increasingly threatened, and in some cases, rather diminished by the various stress factors induced by climate change. Over the centuries, plant scientists have developed many technologies and various innovations aimed at improving and even boosting agricultural productivity. The most recent of all these is plant biotechnology, which facilitates the introduction of desirable traits into plant genomes of interest at the molecular level. Recently, plant biotechnologists have specifically focused onto plant enzymatic molecules known as adenylate cyclases (ACs) in an effort to elucidate how plants do respond and adapt to stressful factors, and for the possible eventual application of such knowledge into the current agricultural context.

Typically and in light of the critical importance of crop plants to humankind, it is thus essential that human populations remain food secure. It has also been noted that achieving food global security whilst reconciling the impacts of the environment on food security is the greatest

challenge presently faced by mankind. Currently, the global rate of growth in the yields of major cereal crops has steadily been declining, from 3.2 percent per year in 1960 to 1.5 percent per year in 2000. The fact is that the actual demand for food is currently not being met, partly because of the environmental stresses due to climate change. Consequently, if the current trends continue, 70 to 100 percent more food will need to be produced by 2050 when the global population would have burgeoned to 9 billion. It is projected that most of this population growth will occur in developing countries, particularly in the sub-Saharan Africa, where food security is already tenuous (Schade and Pimentel, 2010).

Generally, climate change caused by the increases in human greenhouse gas emissions, volcanic and solar effects, have all already significantly contributed towards the global temperature changes that since have been experienced in the past century. As a result of all this, the drought-affected areas in Africa are also projected to increase in extent to cover greater parts of sub-Saharan Africa, thus affecting agriculture and economic growth. Furthermore, since photosynthesis in some important crops is optimal between 20ºC and 25ºC, yields will definitely decline precipitously at temperatures above 30ºC (Dessler, 2012). Studies estimate that the aggregate negative impact of climate change onto the African agricultural output of up to the period 2080-2100 could be between 15 and 30 percent (FAO, 2015). Consequently, climate change is also unfortunately expected to create new opportunistic platforms for the emergence of additional threats such as new or more vigorous crop pests and diseases (Murphy, 2011) and specifically, by expanding their spectral ranges, shortening their disease incubation times and increasing their incidences in water contamination (Dessler, 2012). As a result of all this, the current biological control agents against pests and diseases may then substantially lose their efficacy and the associated costs of weed control systems also dramatically increasing as alien

plants and indigenous weeds do rapidly adapt to the new and ever-changing environmental conditions (Mpandeli *et al*., 2008).

Stress, with respect to plants, is defined as any change in environmental conditions that might reduce or adversely affect the normal growth and/or development of an organism. These stresses can either be biotic or abiotic; and are a particular threat to crop yields in regions that have already been impacted by climate change. The principal biotic stresses come normally from pests, diseases, other plants and herbivores while abiotic stresses include the thermal conditions (heat and cold), water availability (droughts and floods), mineral and nutrient availability (for example, lack of nitrate or excess salt) and also the man-induced pollutants (for example, heavy metals, ozone and acid rain) (Murphy, 2011).

The perception of stress involves an array of multiple signaling pathways, each responding to different forms of stress while also communicating with other pathways to generate a coordinated response by the plant (Murphy, 2011). Plants, typically, can take limited forms of evasive actions in order to avoid a localized stress, for example, by growing in a different direction away from an area of nutrient-poor or polluted soil. Plants can also avoid long-term stresses by special adaptive mechanisms such as the closure of stomata during the day to avoid water losses. Thus as the direction and pace of environmental change becomes clearer, the possibility of taking some prophylactic steps to either mitigate and/or adapt to these changes becomes more realistic (Newton and Edwards, 2007). One such step would be to develop crop plants with effective stress tolerances, that is, plants that are capable of early detection of a potential threat and the timely induction of an appropriate adaptive and/or defensive physiological response (Murphy, 2011).

Since the inception of agriculture onto this planet, mankind has made numerous efforts to develop consequential innovations and technologies aimed at improving agricultural productivity. One such technology, biotechnology, emerged in America in the late 20th century (Murphy, 2011; Carolan, 2012) and it is a technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use (UN Convention on Biological Diversity, 1992). Throughout history, plant breeders have tirelessly sought to genetically modify food crops to improve yield and increased resistance to stress factors. Apparently, while traditional plant breeding methods have tremendously increased the yields of corn and wheat by approximately 100 percent in the last half of the 20th century, these methods are generally very slow and unpredictable (Yadav and Tyagi, 2006). The 20th century focus by plant breeders onto the yield and quality traits of crop plants has somewhat led to the loss of some endogenous resistances in a number of commercial varieties even though this deficiency was partly compensated for by the subsequent development of chemical agents or biocides to control biotic stress.

However, the continued use of agrochemicals has long been increasingly questioned on the grounds of financial costs, safety issues and environmental impact (Murphy, 2011). Biotechnology therefore now provides a viable opportunity to overcome some of the traditional limitations of plant breeding by enabling plant molecular biologists to identify and clone specific genes encoding desirable traits. However, biotechnology does not necessarily replace the conventional breeding methods but indeed rather offers a more precise and better oriented approach (Yadav and Tyagi, 2006) that enables breeders to identify, in great detail, the interacting metabolic pathways that enable a plant to express its genetic potentials, and hence learning more on how to exactly regulate the genes that control such pathways (Duvick, 2005).

Plant biotechnology is also considered a 'greener' technology because of its reduced environmental impacts as compared to the conventional agricultural methods (Gavrilescu and Christi, 2005).

It has also been noted that the 21st century is the era of genomics where functional omics shall aid crop improvement (Gurder, 2005) through genetic transformations of crops for enhanced and improved resistances to stress factors (Chawla, 2002). The approach of plant biotechnology in this regard has been to focus onto plant molecules that systematically are known to be involved in the maintenance of homeostasis and response to stress factors (Hussain *et al*., 2010), and in particular, those molecules involved in the cell-cell signaling and transduction processes. For many decades, the role of ACs as key components of the AC cell signaling system in synthesizing the second messenger, cAMP, from ATP has been extensively studied. However, the presence of ACs and their exact physiological roles in higher plant signal transduction systems had for quite some time remained obscure (Gasumov *et al*., 1999; Lomovatskaya *et al*., 2008). Recently, thirteen proteins from *Arabidopsis thaliana* were bioinformatically identified as putative and possible ACs (Table 1.1) (Gehring, 2010) and one of them, the pentatricopeptide protein (At-PPR), has since been experimentally confirmed as a functional AC (Ruzvidzo *et al*., 2013). Besides this, only three other higher plant ACs have so far been experimentally tested and functionally confirmed, that is, the *Zea mays* pollen signaling protein (PSiP) (Moutinho *et al*., 2001), the *Nicotiana benthamiana* AC (NbAC) (Ito *et al*., 2014) and the *Hippeastrum hybridum* AC (HpAC1) (Swiezawska *et al*., 2014). Of all these four molecules, the NbAC and the HpAC1 are the only ones to have been firmly shown to be involved in stress response.

Of the other remaining twelve uncharacterized proteins from *A. thaliana* (Gehring, 2010), one candidate, the clathrin assembly protein (CAP) encoded by the At1g68110 gene, contains a unique ENTH/ANTH/VHS domain (Berardini *et al*., 2004) and has since been implicated in stress response with regard to pest infestation (De Vos *et al*., 2005; De Vos and Jander, 2009) as well as with regard to certain hormonal treatments, abiotic stresses and pathogens (Schmid *et al*., 2005). Hence, this study therefore specifically focused on the functional characterisation of this putative protein candidate, with the aim of possibly applying the acquired understanding of its functionality into the Southern African agricultural context for the ultimate improvement of crop yield and food security.

1.1.2 Problem Statement

The clathrin assembly protein (CAP) from *A. thaliana*, encoded by the At1g68110 gene and containing an ENTH/ANTH/VHS domain has previously been implicated in stress response (De Vos *et al*., 2005; Schmid *et al*., 2005; De Vos and Jander, 2009) and recently been annotated as a putative adenylate cyclase (Gehring, 2010). However, no study to date has yet experimentally demonstrated the ability of this putative protein to generate cAMP from ATP. As such, a need has therefore been identified to experimentally explore the possible catalytic activity of this novel protein candidate and potentially determining its possible functional role(s) in stress response and adaptation mechanisms.

1.1.3 Aim of the Research

This study primarily sought to ascertain whether the CAP is actually capable of generating cAMP from ATP, and if so, to further elucidate its possible involvement in plant stress response and adaptation mechanisms.

1.1.4 Objectives of the Research

The following specific objectives were set in order to properly and fully address the above given research aim:

1. To isolate and clone the annotated Arabidopsis At1g68110-adenylate cyclase-containing gene fragment (AtCAP-AC) into a stable and viable heterologous prokaryotic expression system.
2. To optimise strategies for the expression profiles of the recombinant AtCAP-AC protein.
3. To determine the endogenous adenylate cyclase activity of the recombinant AtCAP-AC protein.
4. To determine the *in vivo* adenylate cyclase activity of the recombinant AtCAP-AC protein.
5. To optimise strategies for the purification regimes of the recombinant AtCAP-AC protein.
6. To further characterise the enzymatic activity of the recombinant AtCAP-AC protein *in vitro*.
7. To elucidate the exact physiological roles of the clathrin assembly protein in plants through bioinformatics.

1.1.5 Significance of the Research

This study is significant in that a complete functional characterisation of the annotated At1g68110 gene would strongly contribute toward a better understanding of the general mechanisms through which plants respond and adapt to stressful environmental conditions. The obtained improved scientific knowledge regarding the genes responsible for stress response and adaptation mechanisms in plants would subsequently and significantly also contribute towards the possible improved integrated management of both biotic and abiotic stressful conditions of most of the agronomically important crops in the Southern African region. In addition, this research would also immensely improve the current status of our academic scholarship and literature on functional adenylate cyclases in higher plants.

1.2 LITERATURE REVIEW

1.2.1 Second Messengers in Signal Transduction

External influences naturally determine what happens next in the target cells of all affected organisms through the process of signal transduction (Gomperts *et al*., 2009), whereby hormones and neurotransmitters, normally acting as first messengers, transmit their signals from the extracellular space to the intracellular area via a receptor, and a membrane-sited enzyme system which then releases the second messenger inside the cell (Newton *et al*., 1999). Signaling pathways therefore act as a connecting link between sensing the environment and the generation of the appropriate physiological and biochemical response by an involved cell (Verma *et al*., 2013). The chain of steps in signal transduction is generally referred to as the signaling cascade because a wide array of enzymes and various molecules involved in the events which follow the initial stimulus amplify the cellular response (Yanson, 2006). In the case of an exposure to

stress, the signaling cascade of a signal transduction system in response to this would essentially lead to differential transcriptional changes, which then make the plant tolerant to stress (Verma *et al*., 2013).

Sitaramayya (2010) asserted that the transcription factors that are involved in plant stress response arguably display a modular principle of interaction which maintains signal specificity, and enables an assessment and integration at the terminal point of a signal transduction. As such, in cellular signal transduction systems, transcription factors are therefore recognized as important nuclear targets (Lalli and Sassone-Corsi, 1994). Lately, new areas of investigation in signal transduction have been opened up by the advent of high-throughput technologies in genomics and proteomics, and this has driven a paradigm shift towards the systems integrated way of doing biological research (Yanson, 2006). Notably, this has actually led to a new focus, particularly on the role of second messengers in the signal transduction systems of stress response.

Typically, second messengers are a key component of intracellular signal transduction pathways in all cellular organisms because they are the major devices by which the binding of ligands to cell surface receptors is converted into a cellular consequence (Sinha and Sprang, 2006; Gomperts *et al*., 2009). They are diffusible, small intracellular signaling molecules that function as signaling intermediates, according to their localized concentrations within a cell, by recruiting and activating cognate enzymes, and thus allowing signal amplification (Davies, 2004; Krauss, 2014). These molecules are formed as a result of the activation of receptors and can either be cytosolic or membrane-localised (Gomperts *et al*, 2009; Krauss, 2014). The most common second messengers are inositol triphosphate, diacylglycerol, guanosine-5′-triphosphate-3′-

diphosphate, calcium ions and cyclic nucleotide monophosphates (cNMPs) (Sinha and Sprang, 2006).

The cNMPs include cyclic adenosine 3′,5′-monophosphate (cAMP), cyclic guanosine 3′,5′-monophosphate (cGMP) and cyclic bis-(3′,5′)-diguanosine monophosphate (c-diGMP). Among the cNMPs, cAMP and cGMP have been found in plants, animals and prokaryotes whereas c-diGMP is ubiquitous among prokaryotes. The cNMPs have been shown to regulate cellular functions, including gene expression, metabolism and ion flux. At the molecular level, cAMP, cGMP and c-diGMP are synthesized *in vivo* by nucleotidyl cyclases; adenylate cyclases (ACs), guanylate cyclases (GCs) and diguanylate cyclases (di-GCs) respectively (Sinha and Sprang, 2006; Kok, 2007).

1.2.2 The Role of cAMP in Cellular Function

Cyclic AMP was the first second messenger to be discovered (Sutherland *et al.*, 1957). It functions as an activator of downstream proteins that possess cAMP binding sites and are allosterically regulated by cAMP (Krauss, 2014). The catalytic action of ACs, which converts ATP to cAMP, results in an increase of the intracellular levels of cAMP, which in turn leads to either stimulation or repression of specific gene expressions (Lalli and Sassone-Corsi, 1994). Generally, cAMP elicits a response in two main ways. In one way, it binds to a cAMP-dependent protein kinase which then gains the capability to phosphorylate a wide range of other proteins, leading to an immediate cellular response. In the other way, an intra-nuclear effect is exerted through a regulation of the expression of some cAMP-regulated genes via transcription factors known as the cAMP-response element-binding (CREB) proteins, leading to the *de novo* synthesis of proteins which then provoke the cellular response (Newton *et al.*, 1999).

Cyclic AMP has been established as a signaling molecule in both eukaryotes and prokaryotes, including lower plants, but it has always remained elusive in higher plants. Cyclic AMP has also been established as ubiquitous, mediating the action of hormones and neurotransmitters in mammalian and non-mammalian eukaryotic organisms. In prokaryotes however, cAMP acts through catabolite repression, whereby it drives expression of the *lac* operon in the absence of glucose, thus enabling the bacterium to utilize lactose as an alternative energy source (Newton *et al*., 1999).

Generally, cAMP is the product of the catalysis of ATP by ACs. These ACs naturally exist in two forms; the trans-membrane (membrane-bound) and the soluble (intracellular) forms. Trans-membrane ACs (tmACs) are mainly regulated by G-proteins as part of the G-protein coupled receptor pathways and are activated by forskolin, and the fluoride ion while requiring magnesium as the co-factor for activity. Soluble ACs (sACs), on the other hand, are not associated with the plasma membrane and are mainly activated by the bicarbonate and calcium ions, are insensitive to forskolin and the fluoride ion, and generally prefer manganese over magnesium as the co-factor of catalytic activity (Seamon, 1984; Hess *et al*., 2005; Gomperts *et al*., 2009; Rahman *et al*., 2013; Steegborn, 2014). Moreover and in contrast to tmACs, sACs produce cAMP in various intracellular micro-domains, close to specific cAMP targets such as in the nuclei and mitochondria (Ladilov and Appukuttan, 2014). External factors such as the far-red and red light, temperature, exogenous phytohormones, as well as the specific triggering compounds of fungal and bacterial origin, exert a significant influence onto the activities of both plant tmACs and sACs (Lomovatskaya *et al*., 2008).

1.2.3 The Search for Higher Plant Adenylate Cyclases

The existence of cAMP in higher plants (*Embryophyta*) has been doubted and debated (Hintermann and Parish, 1979) because higher plants differ from other eukaryotic organisms in that they lack neurotransmitters and the presence of plant cell walls would seem to compromise the mammalian cAMP/adenylate cyclase model (Newton *et al*., 1999). Additionally, early studies of higher plant ACs were hampered by the fact that the proteins under study were membrane-bound rather than soluble and the attendant presence in crude biological extracts of active cyclic nucleotide phosphodiesterases which hydrolyse cAMP. Furthermore, ACs represent a relatively very low proportion of plant the cellular proteins (Bahzu and Danchin, 1994).

Following the breakthrough discovery of cAMP by Sutherland *et al*. in 1957, numerous studies were then conducted on cAMP and AC across the animal kingdom. While results of the studies on mammals, prokaryotes and lower plants were readily acceptable, early reports on cAMP in higher plants were vehemently criticized. This criticism was mainly based on the fact that the results were merely either presumptive deductions from observed physiological effects of endogenously supplied cAMP or cAMP analogs, or that the conclusions were based solely onto insufficiently rigorous chromatographic identifications.

Traditionally, the demonstration of higher plant AC activity has been conducted through the use of both histochemical and biochemical procedures. On one hand, the histochemical methods were based predominantly onto the standard Wachstein-Meisel lead phosphate precipitation technique, which principally gave early indications of AC activity in *Zea mays*, *Pisum sativum*,

Alnus glutinosa, *Populus* species, *Vicia faba* and *Phaseolus vulgaris*. On the other hand, the biochemical approach suggested the presence of AC activity in *Medicago sativa*, *Spinacea oleracea*, *Ricinis communis*, *Pisum sativum*, *Verticillium albo-atrum* and *Phaseolus vulgaris* (Newton *et al*., 1999). Moreover, mass spectrometry-based analytical techniques also confirmed the presence of cAMP in several plant species but the real strong functional evidence was still lacking (Martinez-Atienza *et al*., 2007).

The procedures discussed above principally revealed that cAMP levels in higher plants were relatively very low, <20 pmol/g fresh weight compared to >250 pmol/g wet weight in animals (Gehring, 2010). Hence, because of the low AC activity and barely detectable amounts of cAMP as well as the questionable experimental procedures used to identify cAMP and/or AC activity, the significance of cAMP in higher plants has generally been doubted and/or overlooked (Ichikawa *et al*., 1997). For this reason, it has been difficult to discern function(s) for cAMP in plants and therefore, the occurrence of ACs has always remained debatable (Katsumata *et al*., 1978; Gomperts *et al*., 2009). Consequently and in 1995, Assmann went on as far as to declare that no plant AC gene had ever been cloned and, moreover in 2006, Linder even asserted that no AC molecule had ever been conclusively identified in higher plants. It has therefore been commonly argued that until a plant AC is either cloned and/or the protein is sufficiently purified to allow for microsequencing and a complete enzymological characterization, the relevance and validity of the AC activity reported in plants would always be open to serious or rather very critical debates (Assmann, 1995).

The era of high-throughput technologies has fortunately now provided methods for assigning function to genes based on their homology, and using the Basic Local Alignment Search Tool

(BLAST). However and according to Wong and Gehring (2013), higher plant ACs could not be identified using BLAST homology searches based on annotated ACs of prokaryotes, lower eukaryotes or animals because they are often part of complex multifunctional proteins with different domains and functions that are not conserved in plants. Hence, BLAST searches with the known and experimentally confirmed ACs from other organisms do not return any plausible candidate molecules. It has also been noted that the Prosite signatures for class I and II ACs ((EYFG[SA]X(2)LWXLYK) and (YRNXW[NS]E[LIVM]RTLHFXG) respectively) are not present in the Arabidopsis proteome even if 2 mismatches are allowed (Gehring, 2010).

Eventually and from 2003 to 2011, six functionally tested GCs from Arabidopsis were positively identified using a 14 amino acid long search motif deduced from an alignment of conserved and functionally assigned amino acids in the catalytic centre (Figure 1.1A) of annotated GCs from lower and higher eukaryotes (Ludidi and Gehring, 2003; Kwezi *et al*., 2007; Meier, 2010; Qi *et al*., 2010; Kwezi *et al*., 2011; Mulaudzi *et al*., 2011). Subsequent to the success of this approach, a 14 amino acid long modified AC core motif was then designed to have substrate specificity for ATP binding (position 3), stabilization of the transition state from ATP to cAMP (positions 12-14), and hydrogen bonding with the adenine (position 1) (Figure 1.1B) (Gehring, 2010).

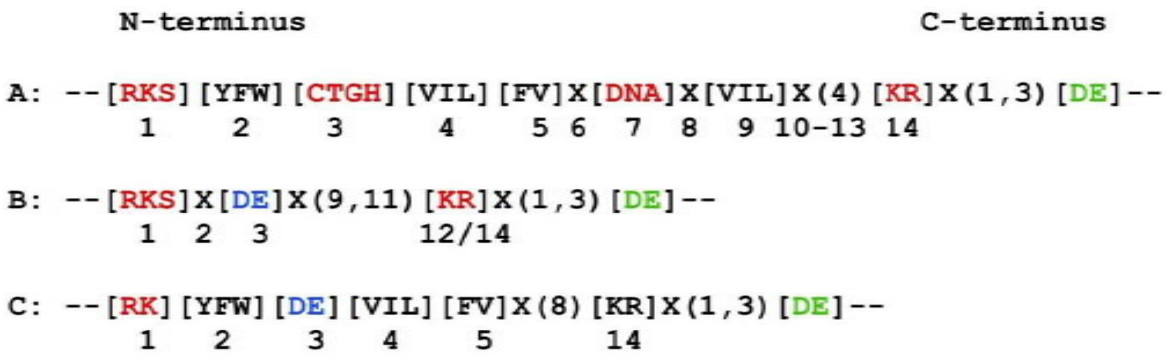

Figure 1.1: The catalytic motifs of nucleotide cyclases. (A) Centre motif of experimentally tested GCs in plants. The residue (red) in position 1 does the hydrogen bonding with the guanine, the amino acid in position 3 confers substrate specificity and the residue in position 14 stabilises the transition (GTP/cGMP). The Mg^{2+}/Mn^{2+}-binding site is C-terminal (green). In the derived motifs (B and C) specific for ACs, position 3 (blue) has been substituted to [DE] to allow for ATP binding (adapted from Gehring, 2013).

A BLAST search of the Arabidopsis genome with this ATP-specific motif the successfully retrieved three annotated but functionally unconfirmed ACs and one unknown protein containing the core (relaxed) AC motif (Table 1.1). However, it was argued that the core motif was not stringent enough to identify candidate ACs *ab initio*. Hence, an extended 16 or 18 amino acid long AC motif was then designed based on the core motif but having the C-terminal Mg^{2+}/Mn^{2+}-binding amino acid (positions 15 or 17) (Figure 1.1C). This extended AC motif then retrieved nine other putative Arabidopsis AC candidates (Table 1.1) (Gehring, 2010).

Table 1.1: The thirteen bioinformatically identified *Arabidopsis thaliana* proteins containing the AC catalytic search motifs (adapted from Gehring, 2010).

ATG NUMBER	SEQUENCE	ANNOTATION
At1g26190*	-SADRVAMRNKNLKR-	Phosphoribulokinase/uridine kinase family protein
At1g73980*	-SVDSRMKYLHGGVSK-	AX4 AC domain containing protein
At2g11890*	-RVEEDEEEIEYWIGK-	G3 AC family protein
At3g21465*	-SSEAKHVENPTEAVK-	Unknown function
At1g25240	-KWEIFEDDFCFTCKDIKE-	Epsin N-terminal homology
At1g62590	-KFDVVISLGEKMQR--LE-	Pentatricopeptide (PPR) protein
At1g68110#	-KWEIFEDDYRCFDR—KD	Clathrin assembly protein
At2g34780	-KFEIVRARNEELKK-EME-	Maternal effect embryo arrest 22
At3g02930	-KFEVVEAGIEAVQR--KE-	Chloroplast protein
At3g04220	-KYDVFPSFRGEDVR--KD-	TIR-NBS-LRR class
At3g18035	-KFDIFQEKVKEIVKVLKD-	Linker histone-like protein – HNO4
At3g28223	-KWEIVSEISPACIKSGLD-	F-box protein
At4g39756	-KWDVVASSFMIERK--CE-	F-box protein

ATG represents the assigned *Arabidopsis thaliana* gene bank numbers for the thirteen putative AC proteins, followed by their amino acid sequences suspected to be their AC catalytic centres, and the names to which each protein was bioinformatically inferred (annotations).

*Proteins that contain the core AC search motif while the rest contain the stringent AC search motif.

#The AtCAP-AC protein that was cloned and functionally characterised in this study.

Subsequent to this notable identification, Ruzvidzo *et al.* (2013) then successfully cloned a gene fragment containing the putative AC catalytic centre of an Arabidopsis pentatricopeptide protein (AtPPR: At1g62590) (one of the genes containing the stringent AC motif designed by Gehring) and firmly showed that the cloned protein fragment could unequivocally produce cAMP from ATP. This was actually demonstrated by its ability to confer AC activity onto the *Escherichia coli* SP850 *cya*A mutant cells, and thus providing non-debatable evidence for its *in vivo* enzymatic activity. Subsequently, this protein was also further characterized and functionally confirmed *in vitro* (unpublished result). Notably, this protein then became the second ever

functional AC to be identified and/or experimentally confirmed in higher plants (Ruzvidzo *et al*., 2013).

Apparently, the first ever higher plant adenylate cyclase to be functionally tested and experimentally confirmed was actually, a pollen signaling protein (PSiP) from *Zea mays* (Moutinho *et al*., 2001). Moutinho *et al*. (2001) too expressed this protein fragment in the SP850 *E. coli cya*A mutant, which eventually resulted in a notable accumulation of cAMP in the transformed bacteria cells, and thus indicating that the PSiP indeed had a *bona fide in vivo* functional AC activity (Moutinho *et al*., 2001).. However, the AC activity of the PSiP has not yet been further functionally characterized *in vitro* (Pietrowska *et al*., 2013). Another higher plant protein, the *Nicotiana benthamiana* adenylate cyclase (NbAC), has recently been shown to possess an *in vitro* and *in vivo* AC activity that is closely associated with cell death during an onset of the wild fire disease, due to an induction by the tabtoxinine-β-lactam bacterial toxin. Here, Ito *et al*. (2014) conducted a database search for putative ACs in tobacco which then clustered with the identified and well-characterised higher plant ACs. Their search retrieved *NbAC* (ACR7530) which showed a 96% amino acid identity with putative ACs from *Zea mays*, *Arabidopsis thaliana* and *Oryza sativa*, including the PSiP and the AtPPR proteins which already had been reported to have functional AC activities, and the At1g25240 and At3g04220 (a clathrin assembly protein and a TIR-NBS-LRR protein from Arabidopsis respectively), which previously had also been retrieved from the Arabidopsis genome by Gehring (2010) using his stringent AC search motif (Ito *et al*., 2014). Nevertheless, the assumption by Ito *et al.* (2014) that there is literally only a single AC in *Nicotiana benthamiana* is highly unlikely considering the contention that no single AC can possibly be responsible for all the known and/or reported cAMP-dependent processes in plants (Gehring, 2010).

Finally and also recently, it has been reported that the *Hippeastrum hybridum* adenylate cyclase (HpAC1) does contain a CYTH-like-Pase domain that is characteristic of bacterial class IV ACs. Apparently, a subsequent BLASTP phylogenetic analysis of this protein then showed that it (HpAC1 protein) indeed does cluster with the putative ACs from *Ipomoea nil*, *Ricinus communis, Zea mays* and *Arabidopsis thaliana*; and also with the other uncharacterized proteins from *Vitis vinifera, Populus trichocarpa* and *Glycine max*. An *in vitro* enzyme activity analysis of this protein with its purified form then showed that it was indeed able to form cAMP from ATP. The *H. hybridum* AC was actually shown to be up-regulated by mechanical damage and the *Phoma narcissi* infection (Swiezawska *et al*., 2014). However, these findings by Swiezawska *et al.* (2014) somewhat raised a number of unanswered questions because they clearly appeared to contain several inconsistencies, such as an assertion that no other plant AC had ever been identified *in vitro* at the time of writing, the lack of an ATP binding site in the HpAC1 sequence, the low homology (51-55%) of the HpAC1 to the other plant ACs which relatively is very low for an AC activity prediction, the absence of an endogenous AC activity due to the HpAC1, and irregularities in the MacConkey agar assaying system for the *cya*A complementation test, among other issues. It is also necessary to point out that the HpAC1 protein had been chemically synthesized and therefore could consequently have had been lacking some of the essential post-translational modifications necessary for its proper enzymatic activities.

Apparently, even though Gehring had bioinformatically identified some putative higher plant ACs in 2010, as recently as 2013, Kast stated that no Arabidopsis gene to that time had ever demonstrated the biological function of AC activity. However, the subsequent enzymological characterization of the various other higher plant proteins as recommended by Assmann (1995)

and accomplished by Moutinho *et al*. (2001), Ruzvidzo *et al.* (2013), Ito *et al*. (2014) and Swiezawska *et al*. (2014) then later clearly confirmed that higher plants (including Arabidopsis) indeed, do possess functional ACs, two of which are involved in responses to environmental stress factors and one in developmental processes.

Ultimately, the discovery of these four higher plant ACs then gives some firm credence to the hypothesis that since cyclic nucleotides have important and diverse roles in plant signaling, it is also highly unlikely that a single AC or just a few can account for all the observed and currently known cAMP-dependent processes in higher plants. In line with this, there is therefore a great potential for the discovery of yet even more higher plant ACs, especially considering that specific cAMP signatures generated in response to biotic and abiotic stress factors act as second messengers in signaling cascades (Gehring, 2010). Apparently, research on the function and modes of action of cAMP in general, has been a significant paradigm shift for the specific area of signal transduction, as it somewhat completely re-shaped and re-modeled the understanding of this specific area (Baker and Kelly, 2004). However, the main challenge now is to conclusively identify and ascertain the exact physiological roles of cAMP in higher plants, particularly by establishing its exact cellular targets in this system. Advantageously, this 'missing link' is now providing us with the actual proof that cAMP is a crucial element in the regulation of basic physiological processes in plants (Newton *et al*., 1999), thus laying the basis for the possible manipulation of cyclic nucleotide systems in higher plants for both the ultimate benefits and sustainability of agriculture and biotechnology. Recently, it has also become obvious that the cyclic nucleotide biochemistry is one of the major areas of research that potentially promises to usher in both the commercial and humanitarian benefits to the world (Newton, 2009).

1.2.4 The Clathrin Assembly Proteins

The main subject of this study was the clathrin assembly protein (CAP, encoded by the At1g68110 gene) from *Arabidopsis thaliana*, which typically contains an ENTH/ANTH/VHS domain, a clathrin adaptor/phosphoinositide binding-GAT-like domain (Berardini *et al*., 2004), an MTH538 TIR-like domain, a fungal immunomodulatory protein (Fve) domain, a threonyl and alanyl tRNA synthetase second additional domain (Kanehisa *et al*., 2014), and a putative adenylate cyclase domain (Gehring, 2010). Naturally, endocytosis is critically an important cellular process for the down-regulation of growth factor signaling, nutrient uptake, antigen presentation, pathogen internalization and maintenance of the plasma membrane (Engqvist-Goldstein *et al*., 2001). During a receptor-mediated endocytosis process, an associated clathrin polymerization process into a polyhedral vesicle coat generally drives a subsequent receptor sorting at the cellular membrane level (Liu *et al*., 1995). Hence, the clathrin-coated membrane systems are typically ubiquitous subcellular structures which are involved in specific intracellular transport processes such as the receptor-mediated endocytosis and the transfer of newly-synthesised proteins from the Golgi apparatus to the lysosomes and secretory storage vesicles (Keen, 1987; Ahle *et al*., 1988; Tebar *et al*., 1999).

According to Morgan *et al*. (2000), clathrin assembly proteins (CAPs) generally mediate clathrin assembly and a connection of the clathrin cage to the budding membranes. The assembly proteins have been surmised to control not only the size of the coated vesicles but also and perhaps more importantly, the selection of membrane proteins for intracellular transport (Ahle and Ungewickell, 1986). In addition, several physiological stimuli are also known to result in increased clathrin coat formations in the living cells (Beck and Keen, 1991). These CAPs naturally exist in various cellular compartments and are usually separated into various

structurally distinct types. Type AP-1 is involved in vesicle trafficking from the *trans*-Golgi apparatus while type AP-2 mediates protein sorting and endocytosis from the plasma membrane. Type AP-3 mediates synaptic vesicle formation from endosomes and protein transport to vacuoles while no function has yet been firmly assigned to the AP-4 type, although its cellular localization suggests a role in *trans*-Golgi transport. Type AP-180 is a synapse-specific protein that is important for synaptic vesicle endocytosis while the clathrin assembly lymphoid myeloid leukemia (CALM) protein is a ubiquitously expressed AP-180 homolog that may also be involved in the formation of clathrin-coated pits. While the exact molecular mechanism of the clathrin assembly protein has not yet been fully clarified, the property of clathrin assembly by the structurally diverse CAPs is strongly believed to be conferred by a conserved DLL motif which is part of a binding site that serves to assemble clathrins by acting as a multivalent cross-linker (Morgan *et al*., 2000).

1.2.5 Functions of the Domains of the CAP Protein

The AP-180 N-terminal homology (ANTH) domain is highly similar to the epsin N-terminal homology (ENTH) domain. Both domains firmly bind inositol phospholipids and proteins, and also markedly contribute towards nucleation and the associated formation of clathrin-coated vesicles (Mao *et al*., 2001). It has been suggested that the E/ANTH domains are actually universal components of the machinery for clathrin-mediated budding since they function with AP-1 and GGA adaptors at the *trans*-Golgi network. All E/ANTH-bearing proteins contain sequences that are indicative of functional roles in clathrin-dependent endocytic events. It has also been suggested that, in addition to recruiting clathrin coat components to membranes, the E/ANTH-bearing proteins probably contribute towards the formation of ordered clathrin coats.

Moreover, a growing body of evidence suggests that the E/ANTH proteins do function as cargo-specific adaptors that recruit endocytic cargos to sites of clathrin-coated vesicle (CCV) formations. Meanwhile, many E/ANTH-bearing proteins have been shown to have significant pools in the cytosol, which then suggests that they have additional functions that are not directly related to the CCV formation. The E/ANTH domains therefore should have other additional protein-binding properties. It has even been suggested that, through the ENTH domain, these proteins may function as transcriptional regulators (Legendre-Guillemin *et al*., 2004).

Structurally, the ENTH domain is similar to the VHS-STAM (Vps27p, hepatocyte growth-regulated tyrosine kinase substrate; Hrs; signal-transducing adaptor molecule) domain that is found at the N-terminus of proteins that participate in membrane trafficking (Legendre-Guillemin *et al*., 2004). The VHS domains interact with the sorting receptors that traffic and transfer cargo between the *trans*-Golgi network and the endosomal compartment, and therefore recognizing the cargo of endocytosis. They typically have a general membrane-targeting and cargo-recognition role in endocytic vesicular transport and signal transduction (Lohi *et al*., 2001; Koshiba *et al*., 2001). Naturally, the Arabidopsis genome encodes for more than 30 ENTH/ANTH/VHS domain-containing proteins but their physiological roles and/or biochemical properties are not well understood because of their generally low sequence conservation that has made functional classification a difficult task (Song *et al*., 2012; Zouhar and Sauer, 2014).

The GAT domain promotes a recruitment of clathrin to the *trans*-Golgi network (Puertollano *et al*., 2001) by binding ubiquitin, which then facilitates the incorporation of cell surface proteins into vesicles (Scott *et al*., 2004). It has also been suggested that the MTH538 domain may have a role in a phosphorylation-independent two-component response regulator system. The

MTH538 domain bears a structural similarity to both the flavodoxin family of proteins and the response regulator proteins of the two-component bacterial signaling pathways (Elcock, 2001). The FIP-fve (Fve) protein is a fruiting body lectin from *Flammulina velutipes* which suppresses allergic symptoms by reducing histamine release and intestinal damage. This Fve domain possesses an immuno-modulatory and anti-inflammatory activity, stimulates lymphocyte mitogenesis, suppresses systemic anaphylaxis reactions and oedema, enhances the transcription of interleukin-2 (IL-2), interferon-γ (IFN-γ) and the tumour necrosis factor-α (TNF-α), and it also haemagglutinates red blood cells and inhibits the lung cancer cell migration (Paaventham *et al*., 2003; Mes and Wichers, 2012; Chang *et al*., 2013). The threonyl and alanyl tRNA synthetase second additional domain (tRNA_SAD) is a dimer which functions in tRNA aminoacylation, ligase activity and ATP binding (Letunic *et al*., 2014) by repressing its own mRNA. This domain also inhibits translational modification by competing with ribosome binding (Sankaranarayanan *et al*., 1999).

From all this information, it is therefore apparent that the AtCAP is a multi-domain, multi-functional protein which possibly, may be involved cAMP-mediated signal transduction pathways by virtue of its possession of the putative AC catalytic domain. Hence, this reported study was indeed premised onto the experimental demonstration of its purported novel AC activity, and the basis for a further elucidation of its possible functional and/or physiological roles in plants.

CHAPTER 2

A PRELIMINARY BIOINFORMATIC ANALYSIS OF THE At1g68110 GENE

ABSTRACT

The clathrin assembly protein (CAP) encoded by the gene At1g68110 in *Arabidopsis thaliana,* has recently been annotated as a possible epsin N-terminal homology (ENTH) protein containing a higher plant adenylate cyclase (AC) domain. Although the protein has been annotated as an ENTH-containing AC molecule, there is still a paucity of information on its experimental and functional studies. In order to create a basis for understanding its possible functional and physiological properties, we employed the various bioinformatic tools and programs such as PHYRE2, PSIPRED, TAIR, NCBI, Genevestigator, ARAMEMNON, KEGG, PSORT II, SUBA3, ATTED-II, ThaleMine, GeneMANIA and STRING to mine the CAP data. Our findings herein indicated that there is literally no current experimental evidence to demonstrate the CAP's *in vivo* translational systems and/or its specific physical and genetic interactions. However and notably, microarray data did implicate this protein in signal transduction and responses to biotic and abiotic stress factors. Thus, our preliminary bioinformatic analysis vividly showed that the annotated CAP is actually a multi-domain, multi-functional protein with a possible role in AC-dependent stress responses and adaptation mechanisms.

2.1 INTRODUCTION

This chapter discusses the current information available on various biological databases about the *Arabidopsis thaliana* gene, At1g68110, and the protein it encodes. As the discussion to follow shall reveal, there is still a meager of information in the current literature related to this At1g68110 gene. Thus, a typical literature review of journals and book texts would surely fail to deliver comprehensive information that may provide a useful picture of the gene and its protein product *visávis* its involvement in both cellular systems and biological processes. For this reason, the tool of bioinformatics analysis was hereby employed to provide an opportunity for the textual mining (information retrieval) of data on this gene/protein in the currently existing biological databases and systems.

Bioinformatics is variously defined as 'the study of information systems in biotic systems' (Hogeweg, 2011); 'conceptualizing biology in terms of macromolecules and then applying "informatics" techniques to understand and organize the information associated with these molecules' (Luscombe *et al*., 2001); and 'the application of computer sciences and allied technologies to answer questions about the mysteries of life' (Nair, 2007). Virtually, it is apparent that the term has evolved from a reference to the study of information processes in biological systems to refer to the creation and use of biological databases. In this regard, an assessment of the relationships among members of a large biological data set can therefore be easily made through the use of algorithms, computational techniques and statistical measures to manage the information contained in databases (Saravanan and Devi, 2012). In this study therefore, knowledge discovery from data (also known as data mining) was conducted using the 'At1g68110' or its amino acid sequence as the query term, to extract 'interesting, non-trivial,

implicit, and potentially useful information' from the data (Wang *et al*., 2005) about this At1g68110 gene/protein.

2.2 DATABASE ANNOTATIONS OF THE AT1G68110 GENE

According to The Arabidopsis Information Resource (TAIR), the ordered locus name, At1g68110, refers to a gene from the eudicot *Arabidopsis thaliana* (L) Heynh., commonly known as the mouse-ear cress or the thale cress (Berardini *et al*., 2004). The gene is listed on various sequence databases, where its RefSeq accession is NM_105481; its NCBI accession number is NP_564922.1; and its UniProt entry is Q9C9X5. At GenPept, it is known as 18408946 and also as LOC843139, whereas its ORF is known as T23K23.4 or T23K23_4 (Thierry-Mieg and Thierry-Mieg, 2006). Its PIR annotation is D96704 and at UniGene, it is At.19366 (Kanehisa, 1997). At the GenBank, it is also known as NC_003070.9 (Benson *et al*., 2013). Among the various genome annotation databases, the gene is known as At1g68110 at the Ensembl Plants (Kersey *et al*., 2013), 843139 at the GeneID (Blanco, Parra and Guigo, 2002) and ath:At1g68110 at the KEGG (Magrane and the UniProt Consortium, 2011). The gene is bioinformatically listed on different databases under various annotations, most probably because the entire Arabidopsis genome has been sequenced while the At1g68110 gene has been picked in the various independent data processing ventures.

2.3 LOCUS DETAIL OF THE AT1G68110 GENE

The At1g68110 gene maps onto chromosome 1, covering about 1.45 kilobases (kb) from position 25528380 to position 25529826 on the direct strand (Thierry-Mieg and Thierry-Mieg, 2006). The At1g68110 gene has one gene model which contains 2 introns and a single exon

(184-1323) (Figure 2.1), having a full length genomic and full length cDNA sequences of 1433 base pairs (bp) apiece, and a full length CDS sequence of 1140 bp (Berardini *et al*., 2004; NCBI, 2015). Hence, we have preliminary evidence here of some post-translational modification of the protein product of this gene but none for the post-transcriptional modification. The genomic DNA is transcribed into a specific protein-coding RNA to proceed to translation, and the mRNA molecule is then recognized by molecules that mediate RNA translation into proteins. The CAP gene is therefore expressed in a protein fusion product.

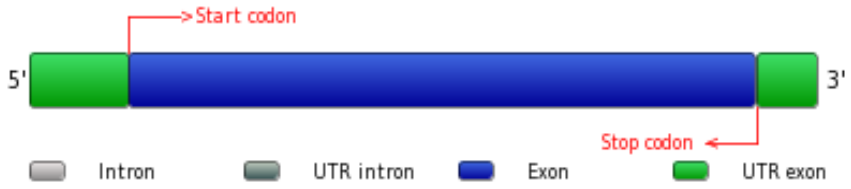

Figure 2.1: The gene model of At1g68110 as presented by the Plaza v2.5 (adapted from Van Bel *et al*., 2012).

2.4 EXPRESSION PROFILE OF THE AT1G68110 GENE

Inferences from expression patterns show that the At1g68110 gene is expressed during 10 growth stages in 15 plant structures, and these growth stages of expression are; 4 anthesis, 4 leaf senescence, C globular, L mature pollen, LP.04 four leaves visible, LP.06 six leaves visible, LP.08 eight leaves visible, LP.10 ten leaves visible, M germinated pollen, and petal differentiation and expansion. The gene is actually expressed in the following plant structures: carpel, collective leaf structure, cotyledon, flower, guard cell, hypocotyl, inflorescence meristem,

leaf apex, petal, plant embryo, pollen, pollen tube cell, sepal, stamen, stem and vascular leaf (Berardini *et al*., 2004).

According to the AceView program, the gene is actually well-expressed at values of up to 0.6 times the average gene in its release. Furthermore, about twenty-one GenBank accessions from 19 cDNA clones provide evidence of the gene's expression in the aerial, flower, leaf, root, silique and flower mixture, bud, inflorescence and hormone-treated callus (Thierry-Mieg and Thierry-Mieg, 2006). According to the Arabidopsis eFP Browser, the gene is highly expressed in the pollen grain (Winter *et al*., 2007). Additionally, data from Genevestigator also shows that the expression of the At1g68110 gene in 98 tissues is highest in 5 tissues as follows (and in a decreasing order of the expression level): pollen, stamen, flower, raceme and petal. Expression is actually highest in the pollen and stamen, and at medium levels in the rest of the tissues (Hruz *et al*., 2008). Furthermore, a developmental map adapted from the Arabidopsis eFP Browser also shows that the gene is highly expressed in the pollen grain relative to the other plant structures (Figure 2.2) (Winter *et al*., 2007). In addition, the AtGenExpress Visualisation Tool (AVT) also shows that the developmental expression of the At1g68110 is highest in flowers, floral organs and young seeds while expression is moderately high in the stem and leaf (Schmid *et al*., 2005).

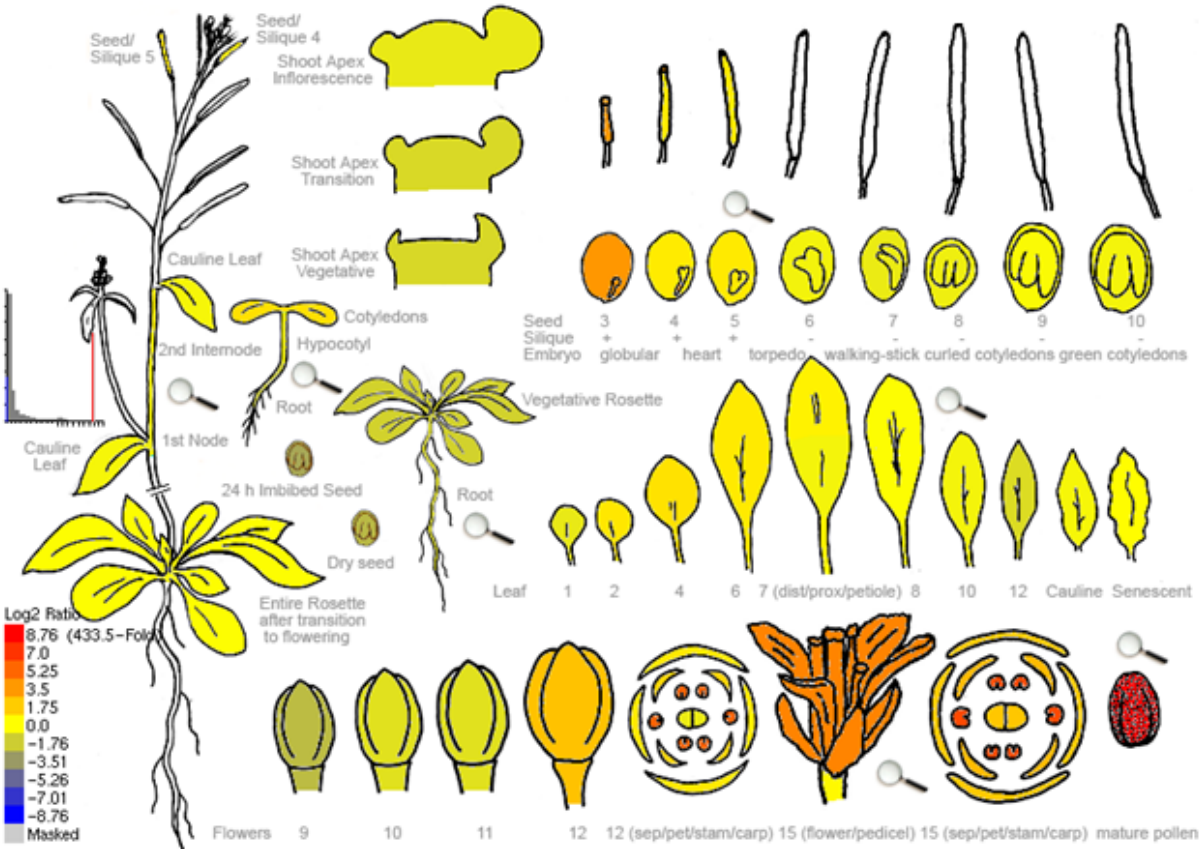

Figure 2.2: The developmental expression profile of the At1g68110 gene (adapted from Winter *et al.*, 2007).

2.5 PROTEIN STRUCTURE OF THE AT1G68110 GENE

According to the TAIR database, the At1g68110 is a protein-coding gene (Berardini *et al.*, 2004) and its protein product has been described as an epsin N-terminal homology (ENTH) domain containing a protein/clathrin assembly protein-related protein (Thierry-Mieg and Thierry-Mieg, 2006). The ARAMEMNON Plant Membrane Protein Database lists the protein as a putative Ap180-like endocytotic accessory candidate (Schwacke *et al.*, 2003). The protein sequence of At1g68110 is, for now, predicted from its mRNA sequence and there is no experimental evidence to date, to demonstrate that this protein is translated *in vivo*. The complete protein

encoded between the first transcription initiation site (atg/Met) and the stop codon is 379 residues (Thierry-Mieg and Thierry-Mieg, 2006). Its calculated molecular weight is 43.8402 kD, and its isoelectric point is 7.8247 (Berardini *et al*., 2004). The BioGRID 3.2 server specifically identifies At1g68110 as a putative clathrin assembly protein with no known physical interactions and no known genetic interactions (Stark *et al*., 2006).

2.5.1 Protein Domains

The protein encoded by the At1g68110 gene belongs to the ENTH/ANTH/VHS superfamily. It contains an epsin-like N-terminal (ENTH) domain, a clathrin adaptor domain, a phosphoinositide-binding domain, a GAT-like domain, an ANTH domain and an ENTH/VHS domain (Berardini*et al*., 2004). An SSDB motif search at the KEGG site revealed that the protein also contains a MTH538 TIR-like domain, a fungal immuno-modulatory protein Fve domain, and a threonyl and alanyl tRNA synthetase second additional domain (Kanehisa *et al*., 2014).

Table 2.1: The domain distribution status in the At1g68110 encoded protein.

DOMAIN	POSITION	PREDICTOR
ANTH	32-294*	Pfam
	165-295*	Superfam
ENTH/VHS	31-155*	Superfam
	33-151*	Superfam
Epsin-like N-terminal	26-158*	Prosite
	30-362*	Superfam
Clathrin adaptor/phosphoinositide binding-GAT-like	172-299*	Superfam
MTH538 TIR-like domain	24-86**	SSDB
Fungal immunomodulatory protein Fve	249-284**	SSDB
Threonyl and alanyl tRNA synthetase second additional domain	251-285**	SSDB
Adenylate cyclase	329-344***	Gehring

*Adapted from Berardini *et al*., 2004.
**Adapted from Kanehisa *et al*., 2014.
***Adapted from Gehring, 2010.

2.5.2 Protein Tertiary Structure

The PHYRE2 server predicts a three-dimensional structure for the At1g68110, and based on the c3zymA protein as a template, with a 100% confidence interval and a 71% coverage (Figure 2.3A) (Kelley and Sternberg, 2009). The 3DLigandSite predicts a binding site composed of residues 143 (Phe), 147 (Tyr), 159 (Arg), 186 (Leu), 190 (Ile), 215 (Glu) and 219 (Ile) (Wass *et al*., 2010). The predicted binding site with a conservation score of >0.5 is shown below in Figure 2.3B in blue colour.

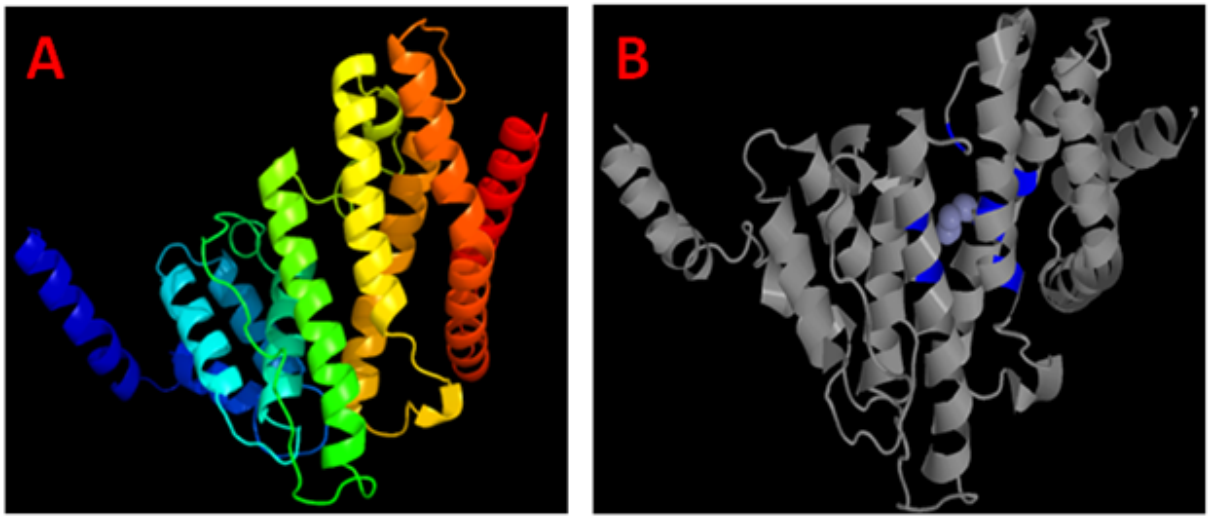

Figure 2.3: The structural model of the AtCAP protein. (A) The predicted tertiary structure of the At1g68110 protein. Rainbow N→C terminus (Kelley and Sternberg, 2009). (B) The predicted ligand binding model of the At1g68110 protein. The binding site is coloured blue (Wass *et al*., 2010).

2.6 PROTEIN FUNCTION OF THE AT1G68110 GENE

The protein product encoded by the At1g68110 gene functions in phospholipid and/or phosphatidylinositol binding, and has been proposed to be involved in the principal process of clathrin coat assembly (Berardini *et al*., 2004). Notably, the PSIPRED program also predicts with high probability (>0.500) that the biological processes in which this gene is involved are transport, regulation of gene expression, RNA processing, protein localization, DNA-dependent regulation of transcription, DNA-dependent transcription, intracellular protein transport, and small GTPase-mediated signal transduction (Buchan *et al*., 2013). Its molecular function is also expected to be phospholipid binding (Berardini *et al*., 2004). The PSIPRED program also predicts with high probability that the other molecular functions of this gene are cytoskeletal protein binding, receptor binding, unfolded protein binding, enzyme regulator activity and DNA binding (Buchan *et al*., 2013).

While the UniProtKB lists endocytosis as a biological process in which the At1g68110 protein is possibly involved (Magrane and the UniProt Consortium, 2011), De Craene *et al*. (2012) revealed that the protein is one of the Picalm9-10 proteins required for the plant endocytic functions. Based on an inference from electronic annotations, the ATTED-II Version 7.1, which provides co-regulated gene relationshipsto estimate gene functions, also further predicts that the At1g68110 protein does bind to 1-phosphatidylinositol (Obayashi *et al*., 2007).

2.7 SUBCELLULAR LOCALIZATION OF THE AT1G68110 PROTEIN

According to TAIR, the protein is located in the mitochondrion (Berardini *et al*., 2004) while AceView proposes that, at organelle level, the protein is located in the clathrin coat (The Gene Ontology Consortium, 2000). Yeast data-trained PSORT II analysis predicts that the subcellular location of this protein is most likely to be the endoplasmic reticulum (33%) or the mitochondria (33%); and less likely in the vacuoles (11%), the Golgi apparatus (11%) or the nucleus (11%) (Nakai and Horton, 1999). On the other hand, the SUBA3 (The SubCellular Proteomic Database) precompiled its bioinformatic subcellular localization predictions as follows: mitochondrion (iPSORT, MitoPred, Mitoprot 2, Predotar, PredSL, PProwler, SLPFA, SLP-Local, TargetP and WoLF PSORT), plasma membrane (AdaBoost), plastid (BaCelLo), nucleus (Plant-mLoc and SubLoc) and cytosol (YLoc) (Heazlewood *et al*., 2013).

2.8 TRANSMEMBRANE TOPOLOGY OF THE AT1G68110 GENE PRODUCT

The PSIPRED server predicts that the At1g68110 encoded protein contains a trans-membrane domain that is located between positions 205 and 220, and as is shown in Figure 2.4 below.

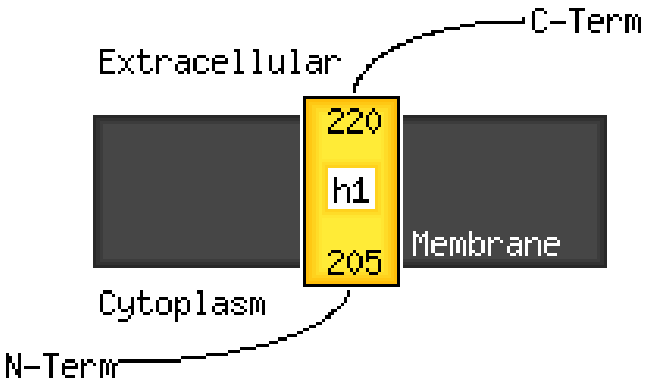

Figure 2.4: The trans-membrane topology of At1g68110 (Buchan *et al*., 2013).

The protein is predicted to have an extracellular region (222-379), a cytoplasmic region (1-205), two disordered protein binding regions (10-11, 157-175), a pore lining helix region (206-221) and three disordered regions (1, 163-164, 298-320) (Buchan *et al*., 2013).

2.9 RECOMBINANT PROTEIN SOLUBILITY OF THE AT1G68110 GENE

Proteins possess ionisable groups andcan be made to exist in solution as electrically charged species. The solubility or lack thereof of a protein has an impact on its physiological and functional properties. Based on its amino acid composition, the At1g68110 recombinant protein is predicted to have a 0.0% chance of solubility when over-expressed in *Escherichia coli* (Diaz *et al*., 2009).

2.10 PROTEIN HOMOLOGY AND GENE CLUSTERING OF THE AT1G68110 GENE

The At1g68110 gene has 367 BLAST hits to 350 proteins in 25 species: Archae - 0; Bacteria - 0; Metazoa - 8; Fungi - 7; Plants - 350; Viruses - 0; Other Eukaryotes - 2 (Berardini *et al*., 2004). The Plant Proteome Database (PPDB) list 12 plant genes (3 from each of the following 4 plants: *A. thaliana*, *Nicotiana benthamiana*, *Oryza sativa* and *Zea mays*) which are related to the At1g68110, and with identities greater than 50%; i.e. At1g25240.1, At1g14686.1, At2g01920.1, NbS00009643g0002.1, NbS00019287g0005.1, NbS00024766g0002.1, Os08g36410.1, Os09g27660.9, Os09g27660.10, GRMZM2G084885_P01, GRMZM2G050556_P01 and GRMZM2G335419_P01 respectively (Sun *et al*., 2009). Additionally, a BLAST P analysis of the At1g68110 protein further reveals some significant alignments with the various *A. thaliana* proteins, including the putative clathrin assembly protein, the ENTH/VHS/GAT family protein, the ENTH.ANTH/VHS superfamily protein, the clathrin coat assembly protein AP-180, the aluminium-activated malate transporter family protein, the leucine-rich repeat trans-membrane protein kinase, the voltage-dependent anion channel 4 protein, the glycosyltransferase family protein and the vesicle-associated membrane 711 protein (Altschul *et al*., 1990).

On one hand, the ThaleMine v1.2.1-4bde138 program lists 13 proteins which are homologous to the At1g68110 protein: WBGene00006751 from *Caenorhabditis elegans*, FBgn0086372 from *Drosophila melanogaster*, ZDB-GENE-100422-6, ZDB-GENE-081104-61, ZDB-GENE-041210-81, ZDB-GENE-040426-761 and ZDB-GENE-030131-6795 from *Danio rerio*, ENSG00000065609 and ENSG00000073921 from *Homo sapiens*, RGD:621054 and

RGD:69276 from *Rattus norvegicus*, and S000001204 and S000003473 from *Saccharomyces cerevisae* (Krishnakumar *et al*., 2014).

On the other hand, the GeneMANIA program also proposes that the At1g68110 protein resides within a cluster of *A. thaliana* proteins with shared domains as is shown in Figure 2.5 below. According to ARAMEMNON, the At1g68110 protein clusters with the following putative AP-180-like endocytic proteins from other species: MELO3C020069 from *Cucumis melo*, Potri.008G140100 from *Populus trichocarpa*, Soly01g005070 and Soly04g008350 from *Solanum lycopersicum*, GSVIVG01011590001 from *Vitis vinifera*, Bradi3g7900 from *Brachypodium distachyon*, GSMUA_Achr7T02390, GSMUA_Achr7T18620 and GSMUA_Achr8T22890 from *Musa acuminate*, LOC_Os08g36410 from *Oryza sativa*; and GRMZM2G050556, GRMZM2G084885 and GRMZM2G335419 from *Zea mays* (Schwacke *et al*., 2003).

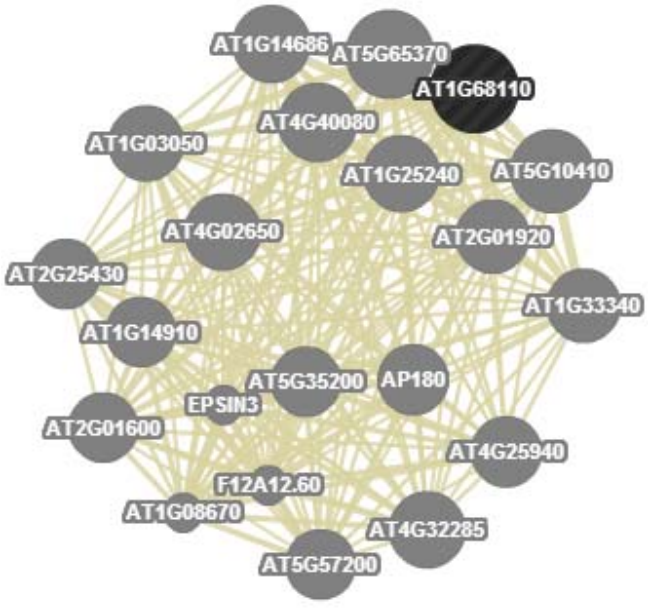

Figure 2.5: A cluster of proteins with domains shared with the At1g68110 (Warde-Farley *et al*., 2010).

2.11 CO-EXPRESSION PATTERNS OF THE AT1G68110 GENE

According to the ATTED-II Version 7.1, At1g68110 is co-expressed with 20 other Arabidopsis genes and is directly connected to At1g14420 (a pectate lyase family protein), At1g78460 (a SOUL heme-binding family protein) and At5g04180 (an alpha carbonic anhydrase 3). The gene is also co-expressed with At2g31500 (CPK24), At2g02720 (lyase), At3g20580 (COBL10), At3g02970 (EXL6), At5g50830, At5g23270 (STPI1), At4g24640 (APPB1), At3g17060 (pectin lyase-like), At5g15110 (lyase), At3g20220 (SAUR-like auxin responsive), At3g05610 (inhibitor), At5g39880, At1g24520 (BCP1), At3g26110 (anther-specific agp 1-like), At4g27300 (kinase) and At2g07180 (kinase) (Obayashi *et al*., 2007).

2.12 PROTEIN-PROTEIN INTERACTIONS OF THE AT1G68110 GENE

According to the STRING 9.1, the predicted functional partners of At1g68110 are At1g21630 and At1g20760 (calcium-binding EF hand family proteins), At3g11130 and At3g08530 (putative clathrin heavy chain proteins), At1g03050 (an ENTH domain-containing protein) and At2g32850 (a protein kinase family protein) (Jensen *et al*., 2009).

2.13 PHYSIOLOGICAL ROLES OF THE AT1G68110 GENE

To date, no phenotype has yet been reported for the At1g68110 gene and furthermore, its *in vivo* function is still unknown (Thierry-Mieg and Thierry-Mieg, 2006). However, an Affymetrix whole-genome microarray to profile the transcriptional regulation in Arabidopsis leaves after exposure to *Myzus persicae* aphid saliva as well as to aphid feeding showed that the At1g68110 is one of the 78 genes whose transcription levels were significantly induced (De Vos *et al*., 2005; De Vos and Jander, 2009). Typically, the gene's expression level increases in response to certain

hormonal treatments such as zeatin; abiotic stresses including osmotic stress, salt stress, drought stress, genotoxic stress, oxidative stress, UV light and wounding in the aerial parts of the plant; pathogens such as Pseudomonas and Phytophthora; and pathogenic elicitors such as $CaCl_2$, $MgCl_2$, GST, HrpZ, GST-LPP1, Flg-22 and LPS. However, the gene shows no response to increased irradiant intensities of UV-light, blue light, red light, far-red light and white light (Schmid *et al*., 2005). Recently, the protein encoded by the At1g68110 gene has also been annotated as a putative adenylate cyclase (Gehring, 2010), implying that it may be involved in signal transduction processes mediated and/or regulated by the second messenger, cAMP.

2.14 CONCLUSION

The protein encoded by the At1g68110 gene is a multi-domain, multi-functional molecule. To date, there is no experimental evidence to demonstrate that this protein product is possibly translated *in vivo*. While there is evidence of some post-translational modifications of this protein, there is still none for the post-transcriptional modification of its mRNA. Additionally, there is also no experimental evidence for any physical and/or genetic interactions in which this gene is involved. However, the gene has been implicated in signal transduction and response to biotic and abiotic stresses.

CHAPTER 3

MOLECULAR CLONING, RECOMBINANT EXPRESSION AND ENDOGENOUS ACTIVITY ASSAYING OF THE AtCAP-AC GENE FRAGMENT FROM *Arabidopsis thaliana*

ABSTRACT

Adenylate cyclases (ACs) are enzymes capable of converting the biological molecule adenosine 5′-triphosphate (ATP) into the product molecule adenosine 3′,5′-cyclic monophosphate (cAMP). Apparently, while cAMP has firmly been established as an important second messenger signaling molecule in higher plants, the existence and/or possible functional roles of ACs in these living systems have always been a subject of serious doubts and critical debates. Recently, a putative clathrin assembly protein encoded by the At1g68110 gene in *Arabidopsis thaliana* (AtCAP) has been bioinformatically annotated as a probable AC molecule and based on a search motif consisting of functionally assigned amino acids in the catalytic centers of annotated and/or experimentally tested nucleotide cyclases. However, even though this putative protein has been bioinformatically annotated as a probable AC candidate, no study to date has ever practically demonstrated its ability to functionally generate cAMP from ATP. In this study therefore, we hereby report the molecular cloning and recombinant expression of an AC-containing fragment domain of this protein molecule (AtCAP-AC) in chemically competent *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cells and practically demonstrate its ability to induce the generation of endogenous cAMP in these prokaryotic expression systems.

3.1 INTRODUCTION

The analysis of gene function usually and at times, always leads to a better comprehension of the physiological processes in living systems (Turcan *et al*., 2011). Currently, greater importance is being attached to the assignment of functions to individual genes in a genome because only a small percentage of annotated genes in databases have any experimental proof for their annotations. Additionally, there is also an advanced drive towards a reduction of reliance onto the nucleotide and/or amino acid sequence similarities as a means of defining gene functions, and thus resulting in the increased use of experimental means and pure proteins to that end (Farrokhi *et al*., 2009).

However, although the use of experimental means and pure proteins is most desirable, the purification process of specific proteins from extracts of biological materials is usually subjective to many difficulties and for many obvious reasons. Firstly, the protein might be present in relatively very low abundances; secondly, the proteins may have multiple isoforms; and/or lastly, the proteins may be found in multiple post-translationally modified forms (Smith, 2005). Apparently, all these specified difficulties can technically be overcome by isolating and expressing the targeted genes of interest in heterologous organisms, and thus allowing for their targeted functional analysis (Frommer and Ninnemann, 1995). Generally, the development of methods through which cDNAs from almost any source can be expressed in suitable host organisms has been facilitated by the relative ease with which cDNAs can now be cloned. This enables the generation of sufficient levels of the targeted proteins for meaningful functional analysis (Seki *et al*., 2002). This heterologous production of proteins principally involves the identification and isolation of a target gene, the identification of a vector and its subsequent

construction into a recombinant system; and the identification and transformation of a suitable host expression system for the recombinant gene, leading to the eventual over-expression of the desired protein molecule, i.e., the directed synthesis of very large amounts of the desired protein (Rai and Padh, 2001).

Among the host organisms for the heterologous expression of proteins, *Escherichia coli* (*E. coli*) is the most widely used bacterial system (Makino *et al*., 2011). It is considered a convenient host for the heterologous expression of foreign proteins because it has relatively high levels of the heterologous gene expression systems, a fast growth rate, the lack of post-translational modifications, an ability to express labeled proteins, and an ease of transformation with exogenous DNA molecules (Pope and Kent, 1996; Francis and Page, 2010). This system is also compatible with almost all commercially available inducible cloning vectors; and its biochemical and genetic information is extensively available (Lueking *et al*., 2000). Literally, recombinant proteins expressed in *E. coli* can comprise up to 50% of its total cellular protein (Francis and Page, 2010). Among the different strains of *E. coli*, the BL21 cells are highly suitable for the heterologous expression of recombinant proteins. These cells are deficient in the *Lon* protease which normally degrades many foreign proteins, and the outer membrane protease (*Omp*T) which normally degrades extracellular proteins (Grodberg and Dunn, 1988). Furthermore, their *hsd*SB mutation always prevents plasmid loss by disruption of the DNA methylation and degradation processes (Rosano and Ceccarelli, 2014).

In the year 2010, Gehring bioinformatically identified a clathrin assembly protein encoded by the At1g68110 gene in *Arabidopsis thaliana* (AtCAP) as a possible adenylate cyclase (AC) candidate, and based on a search motif; [RK][YFW][DE][VIL][FV]X(8)[KR]X(1,3)[DE], which

had specificity for the AC catalytic centre. In this study, we therefore hereby detail the isolation and molecular cloning of a fragment domain of this AtCAP protein candidate which harbours the putative AC catalytic centre (AtCAP-AC), its heterologous expression in the *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cell system, and the determination of its possible endogenous AC catalytic activity through enzyme-immunoassaying system.

3.2 METHODOLOGY

3.2.1 GENERATION AND MAINTENANCE OF ARABIDOPSIS PLANTS

Arabidopsis thaliana ecotype Columbia plants were germinated on Murashige and Skoog (MS) medium and then transplanted into potting soil for growth up to 6 weeks of age under sterile and controlled environmental conditions.

3.2.1.1 Surface Sterilisation of Arabidopsis Seeds

Arabidopsis thaliana seeds (~100) were successively washed (thrice) with 70% (v/v) ethanol in a microfuge tube using a VX-200 vortex mixer (Labnet International Inc., New Jersey, USA) at medium speed. The washing alcohol was discarded followed by another successive wash (thrice) with 500 µl of the seed sterilization buffer (0.1% (w/v) sodium dodecyl sulphate (SDS), 5% (v/v) commercial bleach). Thereafter, the seeds were then successively rinsed (5 times) with 1 ml of sterile distilled water before being suspended in another 1 ml of the sterile distilled water. The washed seeds were then stratified for 3 days at 4ºC (to initiate and promote uniform and synchronized germination).

3.2.1.2 Germination of Arabidopsis Seeds

About 10 stratified seeds were plated onto the sterilized solid MS medium (0.43 (w/v) Murashige and Skoog basal salts, 3% (w/v) sucrose, 0.8% (w/v) type M agar, pH 5.7) under sterile conditions. The seeds were then incubated under sterile growth room conditions (average day/night temperature: 25/16ºC; day/night period: 16/8 hours; light intensity: 10,000 lux) for 14 days in a TL-300 Growth Chamber (Lab Companion Jeio Tech, Seoul, Korea).

3.2.1.3 Growth and Maintenance of Arabidopsis Plants

The 2-week old Arabidopsis seedlings were transplanted into potting soil (40% (v/v) peat-based soil, 40% (v/v) leaf-derived humus and 20% (v/v) vermiculite) and grown under sterile growth room conditions for a further 4 weeks.

3.2.2 EXTRACTION OF TOTAL mRNA FROM ARABIDOPSIS PLANTS

Total mRNA was extracted from the 6-week old *Arabidopsis thaliana* ecotype Columbia leaf tissue using the GeneJET Plant RNA Purification Mini Kit and in accordance with the manufacturer's protocol (Catalogue # K0801; ThermoScientific Inc., Massachusetts, USA).

Briefly, about 100 mg of fresh leaf tissue was flash-frozen in liquid nitrogen followed by its thorough grinding with a pestle and mortar into a fine powder. The tissue powder was then immediately transferred into 500 µl of Plant Lysis Solution in a 1.5 ml microfuge tube and the tube was vortexed at high speed for 20 seconds. The mixture was then incubated at 56ºC for 3 minutes and later centrifuged at 20,000*g* in a LSE High Speed Microcentrifuge (Corning Inc., Amsterdam, Netherlands). The supernatant was transferred into a clean 1.5 ml microfuge tube and mixed with 250 µl of 96% (v/v) ethanol by pipetting. The resultant mixture was then

transferred into a purification column in a collection tube and centrifuged for 1 minute at 12,000*g*. The flow-through was discarded while the purification column and the collection tube were re-assembled. Wash Buffer 1 (700 µl) was added to the purification column and the loaded column was then centrifuged for 1 minute at 12,000*g*. The purification column was placed in a new and clean 1.5 ml collection tube and 500 µl of Wash Buffer 2 was then added to the column. The loaded column was centrifuged for 1 minute at 12,000*g*. After the flow-through was discarded, the wash step was repeated one more time. The washed column was then transferred to an RNAse-free 1.5 ml collection tube and the bound mRNA was then eluted by adding 50 µl of nuclease-free water directly onto the column membrane and centrifuging for 1 minute at 12,000*g*.

The eluted total mRNA was then quantified using a 2000 Nanodrop Spectrophotometer (ThermoScientific Inc., Massachusetts, USA) and resolved on a 1.0% (w/v) agarose gel using a BCMSMINI7 multi-SUB Mini Electrophoresis System (Biocom Ltd, Bridge of Weir, UK) for 40 minutes at 80 volts, and to confirm its purity and integrity. The resolved mRNA was then visualized under ultraviolet (UV) illumination using a UV-2000 Trans-illuminator (Bio-Rad Laboratories, California, USA) and its resultant image captured by a Complete ChemiDoc Imaging System (Bio-Rad Laboratories Inc., California, USA).

3.2.3 PRIMER DESIGNING AND CHEMICAL SYNTHESIS

The genomic sequence of the At1g68110 gene as retrieved from The Arabidopsis Information Resource (TAIR, www.arabidopis.org) revealed that the gene does not undergo any post-transcriptional modification (as evidenced by the fact that its full length genomic and copy DNA (cDNA) sequences are both 1433 base pairs (bp) long). However, the gene appears to be

involved in some form of post-translational modifications because its full length coding DNA sequence (CDS) is 1140 bp. Hence, based on the genetic information previously provided for the At1g68110 gene (Gehring, 2010), two sequence-specific primers (forward and reverse) were then manually designed for it and in such a way that they could fully flank the N- and C- termini ends of its AC-containing catalytic centre (Figure 3.1), and also so that the desired AtCAP-AC gene fragment could then be specifically amplified. The designed primer sequences were then sent over to the Inqaba Biotechnological Sciences (Pretoria, RSA) for chemical synthesis and subsequent supply.

```
         10         20         30         40         50
MKLWKRAAAA IKDRKSLLAV GFSRRNSSYR NADLEAAIIK ATSHDDSSVD
         60         70         80         90        100
YSNAHRVYKW IRSSPLNLKT LVYAISSRVN HTRSWIVALK SLMLLHGVLC
        110        120        130        140        150
CKVPSVVGEF RRLPFDLSDF SDGHSCLSKT WGFNVFVRTY FAFLHHYSSF
        160        170        180        190        200
LSDQIHRLRG NNRRSLEKTS DSVIQELERI QKLQSLLDMI LQIRPVADNM
        210        220        230        240        250
KKTLILEAMD CLVIESINIY GRICGAVMKV LPLAGKSEAA TVLKIVNKTT
        260        270        280        290        300
SQGEDLIVYF EFCKGFGVSN AREIPQFVRI PEEEVEAIEK MIDTVQEKPK
        310        320        330        340        350
LEKDEEKEDE KAMVVLEQPK KLQTIITDKW EIFEDDYRCF DRKDKWEIFE
        360        370
DEYHQNHLPL ITMNQPVYIT YTMPDLITF

At1g68110-AC Forward Primer
5'-gAA TTC TgC AAA ggT TTC ggTgTCTCg AAC-3'
At1g68110-AC Reverse Primer
5'-gAATgT AAT CAA ATC Tgg CAT TgT ATA AgT-3'
```

Figure 3.1: A complete amino acid sequence of the At1g68110 protein showing its annotated AC catalytic centre (red highlight and bold) and the priming sites of this targeted catalytic center (yellow highlights and bold). Nucleotide sequences of both the forward and reverse primers for this targeted catalytic center are also presented (blue highlights).

3.2.4 ISOLATION AND AMPLIFICATION OF THE AtCAP-AC GENE FRAGMENT

The targeted AtCAP-AC gene fragment was isolated from the total mRNA previously extracted from leaves of the 6-week old Arabidopsis plants using a Verso 1-Step RT-PCR Reddy Mix System (Catalogue # AB-1454/LD/A; Thermo Scientific Inc., California, USA), whereby the extracted total mRNA was first used as a template to generate cDNA and then the generated cDNA (together with the two acquired sequence-specific primers) as a template to amplify the targeted AtCAP-AC gene fragment. An RT-PCR reaction mix of 50 µl was prepared as is shown in Table 3.1 below.

Table 3.1: Components of the RT-PCR reaction mix for the targeted isolation of the desired AtCAP-AC gene fragment.

Component	Volume (µl)	Final Concentration
Verso Enzyme Mix	1	
1-Step RT-PCR Reddy Master Mix	25	1X
Forward Primer	1	200 nM
Reverse Primer	1	200 nM
RT Enhancer	2.5	
PCR Grade Water	19.5	
RNA Template	1	1 ng

The thermal cycling system for the amplification of the targeted AtCAP-AC gene fragment was then conducted on a C1000 Touch Thermal Cycler (BioRad Laboratories, California, USA), whose cycling conditions are provided below in Table 3.2.

Table 3.2: Thermal cycling conditions used for the isolation of the AtCAP-AC gene fragment via a 1-step RT-PCR system.

Step	Temperature (ºC)	Time (mins)	Number of Cycles
cDNA Synthesis	50	15.00	1
Verso Inactivation	95	2.00	1
Denaturation	95	0.33	45
Annealing	60	0.50	
Extension	72	1.00	
Final Extension	72	5.00	1

The amplified AtCAP-AC gene fragment was then resolved on a 1.0% (w/v) agarose gel supplemented with 0.5 µg/ml ethidium bromide, whereby the amplicon was resolved against a 100 bp Gene-Ruler DNA ladder in 1X TBE buffer for 50 minutes at 80 volts. The resolved gel was then visualized under UV light using a 2000 UV Trans-illuminator (Bio-Rad Laboratories Inc., California, USA) and images subsequently captured with a Complete ChemiDoc Imaging System (Bio-Rad Laboratories Inc, California, USA).

3.2.5 MOLECULAR CLONING OF THE AtCAP-AC GENE FRAGMENT

The isolated AtCAP-AC gene fragment was ligated into the pTrcHis2-TOPO expression vector with the resultant pTrcHis2-TOPO:AtCAP-AC fusion construct maintained in chemically competent One Shot TOPO 10 *Escherichia coli* cells and in accordance with the pTrcHis2-TOPO TA Expression protocol (Catalogue # K4400-40; Invitrogen Corp., New York, USA).

3.2.5.1 Post-amplification Addition of the 3′-adenine Overhangs to the AtCAP-AC Gene Fragment

In order to facilitate and increase the cloning efficiency of the AtCAP-AC insert into the pTrcHis2-TOPO expression vector, single A-overhangs were added to the 3′ ends of the blunt-ended AtCAP-AC gene fragment. About 1 unit (1 µl) of *Taq* polymerase was added to the RT-PCR product (40 µl) on ice and the mixture incubated for 10 minutes at 72ºC on a C1000 Touch Thermal Cycler (Bio-Rad Laboratories, California, USA). The reaction tube was then placed on ice subsequent to its use in the pTrcHis2-TOPO TA Expression cloning system.

3.2.5.2 Ligation of the AtCAP-AC Gene Fragment into the pTrcHis2-TOPO Expression Vector

An aliquot (4 µl) of the adenylated AtCAP-AC gene fragment was added to 1 µl (10 ng) of the ice-cold pTrcHis2-TOPO expression vector and the mixture was then gently mixed with a pipette before its incubation at room temperature for 5 minutes. The resultant ligation mixture was later used for the transformation of chemically competent One Shot TOPO 10 *E. coli* expression cells.

3.2.5.3 Transformation of the Chemically Competent *E. coli* One Shot TOPO 10 Cells with the pTrcHis2-TOPO:AtCAP-AC Expression Construct

Immediately after the ligation process, 2 µl of the ligation mixture (pTrcHis2-TOPO:AtCAP-AC fusion construct) were added to an ice-cold vial containing 40 µl of the chemically competent One Shot TOPO 10 *E. coli* cells (Invitrogen Corp., New York, USA) and the mixture was gently mixed before being incubated on ice for 30 minutes. The mixture was then heat-shocked at 42ºC for 30 seconds on an AccuBlock Digital Dry Bath (Labnet International Inc., New Jersey, USA).

The tube was immediately transferred onto ice and 250 µl of room temperature SOC broth (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM sodium chloride, 2.5 mM potassium chloride, 10 mM magnesium chloride, 10 mM magnesium sulphate, 20 mM glucose) was added. The culture was then shaken horizontally in an SI-600 Bench-top Shaking Incubator (Lab Companion GMI Inc., Minnesota, USA) for 30 minutes at 200 rpm at 37ºC. About 50 µl of the transformation culture was then spread-plated onto a pre-warmed selective LB agar plate (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, 0.5% (w/v) glucose, 1.5% (w/v) agar, 50 µg/ml ampicillin, pH: 7.0) and incubated overnight at 37ºC in an Incubat 2001651 Table-top Incubator (JP Selecta SA., Barcelona, Spain). The plates were then visually analysed for colony growth in the next morning.

3.2.5.4 Screening of Positive Clones

A rigorous screening procedure was undertaken to, firstly, check if the ligation process was successful, secondly, to determine if such a successful ligation process was in the correct expression orientation and lastly, to check if the ligated fragment was in the correct and acceptable reading frame.

3.2.5.4.1 Extraction of the Plasmid

From the selective LB plate, several single colonies were selected and each used to inoculate individual 10 ml selective LB broth media (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) sodium chloride, 0.5% (w/v) glucose, 50 µg/ml ampicillin, pH: 7.0). The cultures were then incubated overnight at 37ºC in a shaking incubator at 200 rpm. Cells from the overnight cultures were then harvested by centrifugation at 4,000*g* using a TD3 800B Low Speed Table-

top Centrifuge (Xiangyi Laboratory Instrument Development Ltd., Hunan, China) for 20 minutes. The supernatant was discarded while plasmid extraction was then undertaken from the harvested cells using the GeneJET Plasmid Miniprep Kit and in accordance with the manufacturer's protocol and instructions (Catalogue # K0503; ThermoScientific Inc., Massachusetts, USA).

Briefly, the harvested cells were re-suspended into 250 µl of the Resuspension Solution. A total volume of 250 µl of the Lysis Solution was also added to the cell suspension and the mixture was then gently mixed by inversion. An aliquot of about 350 µl of the Neutralisation Solution was also added to the cell suspension and the mixture, once again, was gently mixed by inversion. The cell suspension was then centrifuged at 16,300*g* for 5 minutes before the supernatant was transferred to a clean GeneJET Spin Column and further centrifuged at 16,300*g* for 1 minute. The flow-through was discarded. The column was washed using 500 µl of the Wash Solution at 16,300*g* for 1 minute and the flow-through was discarded. The wash step was repeated one more time, with an additional 1 minute centrifugation to remove any residual Wash Solution and/or chemical elements. The washed column was transferred to a clean 1.5 ml microfuge tube and 50 µl of pre-warmed (56ºC) water was then added directly onto its membrane followed by incubation at room temperature for 2 minutes. The assembled column was then centrifuged at 16,300*g* for 1 minute to elute the plasmid. The eluted plamid was then collected and kept in the 1.5 microfuge while the used column was dessembled and discarded. The DNA concentration of the eluted plasmid (fusion construct) was then quantified using a 2000 Nanodrop Spectrophotometer (ThermoScientific Inc., Massachusetts, USA).

3.2.5.4.2 Analysis of the positive constructs

Analysis of the positive constructs was carried out through a normal PCR procedure and in accordance with the standard MyTaq Mix PCR protocol (Catalogue # BIO-25041; Bioline, London, UK) firstly, to check if the AtCAP-AC gene fragment insert was successfully ligated into the pTrcHis2-TOPO expression vector and secondly, to confirm if such a ligation was in the correct orientation. The reaction mixtures for these two PCR processes are respectively shown below in Tables 3.3 and 3.5, while their associated thermal cycling conditions are also shown in Table 3.4 and 3.6 respectively.

Table 3.3: Reaction components of a PCR reaction mixture to confirm the successful cloning of the AtCAP-AC gene fragment insert into the pTrcHis2-TOPO expression vector.

Component	**Volume (µl)**	**Final Concentration**
Plasmid DNA Template	1	2 ng
Insert Forward Primer	1	200 nM
Insert Reverse Primer	1	200 nM
MyTaq Reaction Mix	25	1X
PCR Grade Water	22	

Table 3.4: The thermal cycling conditions for confirmation of the successful cloning of the AtCAP-AC gene fragment insert into the pTrcHis2-TOPO expression vector.

Step	**Temperature (ºC)**	**Time (mins)**	**Cycles**
Initial Denaturation	95	1.00	1
Denaturing	95	0.25	35
Annealing	60	0.25	
Extension	72	10.00	

Table 3.5: Reaction components of a PCR reaction mixture to confirm the correct orientation of the AtCAP-AC gene insert in the pTrcHis2-TOPO expression vector.

Component	Volume (µl)	Final Concentration
Plasmid DNA Template	1	2 ng
Insert Forward Primer	1	200 nM
Vector Reverse Primer	1	200 nM
MyTaq Reaction Mix	25	1X
PCR Grade Water	22	

Table 3.6: The thermal cycling conditions for the confirmation of the correct orientation of the AtCAP-AC gene insert in the pTrcHis2-TOPO expression vector.

Step	Temperature (ºC)	Time (mins)	Cycles
Initial Denaturation	95	1.00	1
Denaturing	95	0.25	35
Annealing	50	0.25	
Extension	72	10.00	

The PCR products from Tables 3.3 and 3.5 were resolved on a 1% (w/v) agarose gel supplemented with 0.5 µg/ml ethidium bromide and submerged in 1X TBE buffer for 50 minutes at 80 volts. The resolved gel was then visualized under UV light using a 2000 UV Trans-illuminator (Bio-Rad Laboratories Inc., California, USA) and images subsequently captured by a Complete ChemiDoc Imaging System (Bio-Rad Laboratories Inc., California, USA). A double amplification of the targeted AtCAP-AC gene insert in both reaction samples of Tables 3.3 and 3.5 was therefore, correspondingly confirming its successful cloning and correct orientation in the pTrcHis2-TOPO expression vector. In addition, all the successfully confirmed clones were further sent over to the Inqaba Biotechnological Sciences (Pretoria, RSA) for a validation of their insert in-frame orientations within the pTrcHis2-TOPO expression vector through sequencing.

3.2.5.5 Transformation of the *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs Cells with the pTrcHis2-TOPO:AtCAP-AC Expression Construct

After confirming a particular clone to be positive in section 3.2.5.4.2 above, its respective pTrcHis2-TOPO:AtCAP-AC fusion construct was then used to transform some chemically competent *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cells and in accordance with the manufacturer's instructions (Catalogue # MA019; Lucigen Corp., Wisconsin, USA), and in preparation for the subsequent expression of the desired recombinant AtCAP-AC fusion product.

Briefly, a total amount of 200 ng (1 µl) of the ice-cold pTrcHis2-TOPO:AtCAP-AC fusion construct was gently mixed with 100 µl of ice-cold chemically competent *E. cloni* BL21 (DE3) pLysS DUOs cells and the mixture further incubated on ice for 20 minutes. The mixture was then heat-shocked at 42ºC on a heating block for 90 seconds before an immediate re-incubation on ice for 5 minutes. The mixture was then mixed with 500 µl of the SOC broth and incubated at 37ºC for 90 minutes at 200 rpm. About 50 µl of the cell culture was then spread-plated onto a selective LB agar plate (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) sodium chloride, 0.5% (w/v) glucose, 1.0% (w/v) agar, 100 µg/ml ampicillin, 34 µg/ml chloramphenicol, pH: 7.0). The plate was then incubated at 37ºC overnight and later visually analysed for colony growth in the next morning.

3.2.6 RECOMBINANT EXPRESSION AND ENDOGENOUS ACTIVITY ASSAYING OF THE AtCAP-AC PROTEIN

The AtCAP-AC protein product was recombinantly expressed in chemically competent *E. cloni* BL21 (DE3) pLysS DUOs cells followed by an assessment of its possible endogenous AC activity through enzyme immunoassay.

3.2.6.1 Recombinant Expression of the AtCAP-AC Protein

A single colony of the *E. cloni* BL21 (DE3) pLysS DUOs cells harbouring the pTrcHis2-TOPO:AtCAP-AC fusion construct was used to inoculate 10 ml of the selective LB broth (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) sodium chloride, 0.5% (w/v) glucose, 100 µg/ml ampicillin, 34 µg/ml chloramphenicol, pH: 7.0) and the culture incubated at 37ºC overnight at 200 rpm. The next morning, 1 ml of the overnight culture was then used to inoculate 20 ml of the selective LB broth. The culture was incubated at 37ºC at 200 rpm and up until its optical density (OD_{600}) had reached 0.5, as was measured by the Spectronic Helios Epsilon Spectrophotometer (Thermo Electronic Scientific Instruments LLC., Wisconsin, USA). At this desired optical density, the culture was immediately chilled on ice and divided into four separate portions of 5 ml each.

Protein expression was then induced in three of the four cultures by adding 1 mM IPTG (Sigma-Aldrich Corp., Missouri, USA) while the other tube was left un-induced (control). For two of the three induced cultures, one was further supplemented with 100 µM of the AC inducer, forskolin (Sigma-Aldrich Corp., Missouri, USA), while the other was equally supplemented with 100 µM of the potent AC inhibitor, 2′,5′-dideoxyadenosine (Sigma-Aldrich Corp., Missouri, USA). The four separate cell cultures were then allowed to grow for a further 1 hour at 37ºC at 200 rpm. At the end of the 1 hour, 500 µl of each cell culture was separately collected for analysis by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE), while the rest of the cells were also separately harvested by centrifugation at 16,300*g* for 5 minutes for further analysis by enzyme-immunoassaying.

3.2.6.2 Determination of the Endogenous AC Activity of the Recombinant AtCAP-AC Protein

The harvested cell cultures from section 3.2.6.1 above were each resuspended in 1 ml Lysis Buffer 1 (Amersham Healthcare, New Jersey, USA) supplemented with 2 mM 3-isobutyl-1-methylxanthine (IBMX) to inhibit phosphodiesterases. The resuspended cells were subsequently lysed through shaking on an orbital shaker at 200 rpm for 1 hour at room temperature. The generated crude lysate was then clarified into a clear supernatant through centrifugation at 16,300*g* for 5 minutes. About 200 µl of this cleared supernatant were then thoroughly mixed with 200 µl of Lysis Buffer 2 (Amersham Healthcare, New Jersey, USA). The resultant solution was then acetylated by adding the acetylating reagent (2:1 triethylamine:acetic anhydride (v/v)) at a volumetric ratio of 1:20 acetylating reagent:sample (v/v). The mixture was then thoroughly mixed by vortexing at high speed for 2 seconds, and before the inherent cAMP levels in each of the prepared solutions were measured in triplicate sets by the cAMP-linked enzyme-immunoassaying system and in accordance with the manufacturer's protocol (Catalogue # CA201; Sigma-Aldrich Corp., Missouri, USA).

3.2.7 STATISTICAL ANALYSIS OF THE ENZYME-IMMUNOASSAY DATA

The results of the immunoassaying system obtained here were all based on the means of three replicates where outcomes of each assay were sequentially subjected to analysis of variance (ANOVA) (Super-Anova, Statsgraphics Version 7, 1993, Statsgraphics Corporation, USA). Notably, wherever the ANOVA revealed some significant differences between outcomes, the means (n = 3) were then separated by the *post hoc* Student Newman Kuehls (SNK), multiple range test ($p \leq 0.05$).

3.3 RESULTS

3.3.1 Isolation of the AtCAP-AC Gene Fragment from Arabidopsis

Following the bioinformatic annotation of the At1g68110 gene as a possible AC molecule (Gehring, 2010), an attempt was therefore made in this study to isolate and possibly further characterize its truncated fragment domain harbouring the AC catalytic center (AtCAP-AC). Using a specialized 1-step RT-PCR technique (ThermoScientific Inc., Massachusetts,USA), the targeted AtCAP-AC gene fragment was successfully isolated (Figure 3.2 below) and the same truncated fragment was then later on used for the further and subsequent biochemical and functional characterization of the At1g68110 gene in the study.

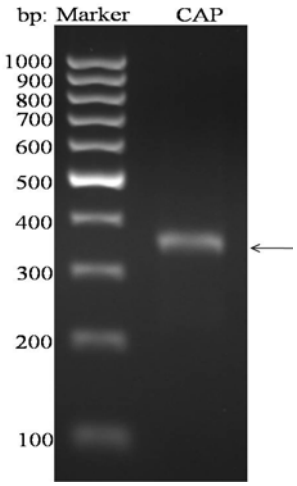

Figure 3.2: Isolation of the AtCAP-AC gene fragment from *Arabidopsis thaliana*. The AtCAP-AC gene fragment was isolated from the total Arabidopsis mRNA using two sequence-specific primers via a specialized 1-step RT-PCR system. The amplified AtCAP-AC gene fragment was then resolved by electrophoresis on a 1.0% (w/v) agarose gel at 80 volts for 45 minutes. Marker represents a 100 bp GeneRuler ladder (Thermo Scientific Inc., Burlington, Canada), CAP represents the amplified AtCAP-AC gene fragment, while the arrow marks the resolved 360 bp band of the amplified AtCAP-AC gene fragment.

3.3.2 Cloning of the AtCAP-AC Gene Fragment

After the successful isolation of the targeted AtCAP-AC gene fragment from the Arabidopsis plant, this gene fragment was then ligated into a pTrcHis2-TOPO expression vector to yield a pTrcHis2-TOPO:AtCAP-AC fusion construct. The pTrcHis2-TOPO:AtCAP-AC fusion construct was then assessed to ascertain whether the ligated AtCAP-AC gene fragment was cloned in the correct orientation (Figure 3.3). This step was crucial before the AtCAP-AC gene fragment could be expressed into its respective and most desired recombinant protein.

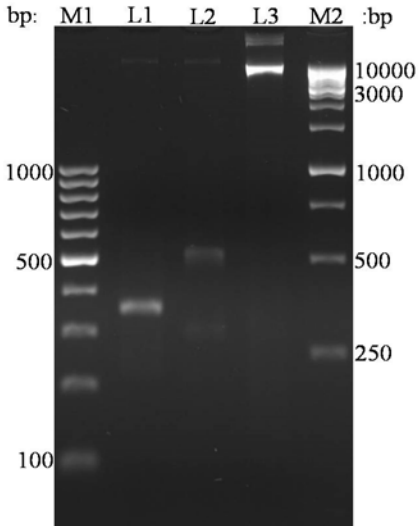

Figure 3.3: Screening of the correct pTrcHis2-TOPO:AtCAP-AC fusion construct. The cloned AtCAP-AC gene fragment was first assessed for its correct orientation within the pTrcHis2-TOPO expression vector before its subsequent use for the expression of the desired recombinant AtCAP-AC protein. This was done by re-amplifying it with its own primers as well as with one of its own primers and one of the pTrcHis2-TOPO vector primers. All amplified PCR products were then resolved on a 1.0% (w/v) agarose gel, where M1 is the 100 bp GeneRuler ladder (Thermo Scientific Inc., Burlington, Canada), L1 is the AtCAP-AC gene fragment amplified with its own primers, L2 is the AtCAP-AC gene fragment amplified with its own forward primer and the pTrcHis2-TOPO vector reverse primer, L3 is the complete pTrcHis2-TOPO:AtCAP-AC fusion expression construct while M2 is the 1 kb GeneRuler ladder (Thermo Scientific, Burlington, Canada).

3.3.3 Recombinant Expression of the AtCAP-AC Protein

In order to facilitate expression of the desired recombinant AtCAP-AC protein, the correctly confirmed pTrcHis2-TOPO:AtCAP-AC fusion construct was used to transform some chemically competent *E. cloni* BL21 (DE3) pLysS DUOs cells followed by the induction of protein expression with 1 mM IPTG. As is shown in Figure 3.4, the desired recombinant AtCAP-AC protein was successfully expressed and also as a His-tagged (C-terminus) fusion product (Crowe *et al*., 1994).

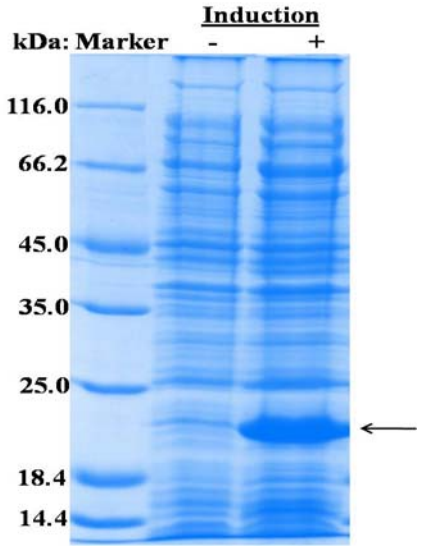

Figure 3.4: Recombinant expression of the AtCAP-AC protein. The recombinant protein fragment was expressed in chemically competent *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cells through their induction with 1 mM IPTG. The expressed protein was then resolved by SDS-PAGE on a 12% poly-acrylamide gel, where Marker is the unstained protein molecular weight marker (ThermoFisher Scientific Inc., New York, USA), while (-) and (+) respectively represent the un-induced and induced cell cultures. The arrow marks the expressed recombinant AtCAP-AC protein.

3.3.4 Endogenous Adenylate Cyclase Activity of the Recombinant AtCAP-AC Protein

In order to determine the endogenous adenylate cyclase activity of the recombinant AtCAP-AC protein, its ability to possibly generate or induce the generation of cAMP within the transformed *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cells under various growth conditions, was assessed and determined by the cAMP-linked enzyme immunoassaying system (Catalogue # CA201; Sigma-Aldrich, Missouri, USA) and as is shown below in Figure 3.5.

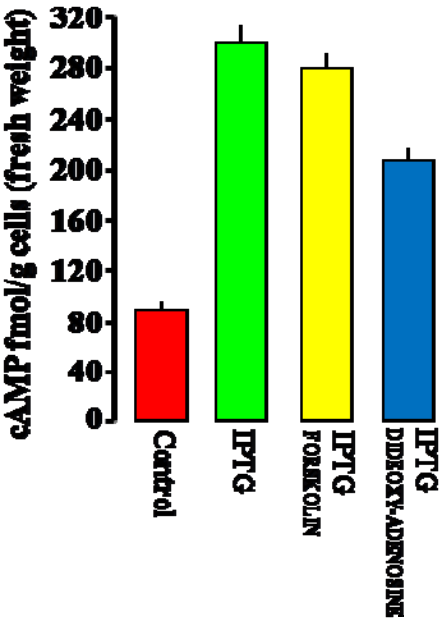

Figure 3.5: Determination of the endogenous adenylate cyclase activity of the recombinant AtCAP-AC protein. Cyclic AMP levels generated by the un-induced (Control) and induced (IPTG) *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cells harbouring the AtCAP-AC gene fragment, and also from the induced cell cultures in the presence of either forskolin or dideoxyadenosine. All cAMP levels were determined by the cAMP-linked enzyme immunoassaying system based on its acetylation protocol (Catalogue # CA201; Sigma-Aldrich Inc., Missouri, USA), where error bars represent the standard errors of the mean readings (n = 3).

3.4 DISCUSSION

In order to facilitate the heterologous expression of the recombinant AtCAP-AC, the expected 360 bp AtCAP-AC gene fragment was first isolated from the leaf tissues of the higher plant, *Arabidopsis thaliana,* via a reverse transcriptase polymerase chain reaction (RT-PCR) system (Figure 3.2). Apparently, plant tissue generally presents specific problems for protein extraction because; firstly, the rigid cellulosic wall must be sheared off to release the cell contents; secondly, the presence of contaminating compounds may cause protein degradation, modification and/or loss of activity; and thirdly, the low levels of protein may exist as part of highly complex protein mixtures (Fido *et al*., 2004). Alternatively, the gene that encodes the targeted and/or desired plant protein can be extracted and expressed in bacterial, fungal or animal cell systems; however, the isolation of genomic DNA from eukaryotic organisms for protein expression may still result in a non-functional protein product due to the presence of introns.

Notably, this situation can be practically circumvented by expressing those desired plant genes in fungal and/or animal cells which, as a result of their eukaryotic nature, are able to process the required post-transcriptional modifications required for the functionality of the targeted recombinant proteins (Clark and Pazdernik, 2011). However, eukaryotic systems still present a big challenge in terms of the complexity of their physiological systems, which in turn makes the aspect of protein processing a huge problem. Apparently, this challenge can still be further overcomed by cloning the copy DNA (cDNA) of the plant gene in a bacterial system because this type of DNA lacks introns and consists solely of the uninterrupted coding sequences. Furthermore, the bacterial system can also be induced to over-express the desired recombinant plant protein (Clark and Pazdernik, 2011). For these reasons and in order for us to express a

functional recombinant AtCAP-AC protein in this study, we therefore utilized the RT-PCR technique to synthesise the cDNA from mRNA and then isolated the targeted AtCAP-AC gene fragment, with an objective of heterologously expressing it in a bacterial system.

The isolated AtCAP-AC gene fragment was then ligated into a pTrcHis2-TOPO plasmid vector that utilizes the TA cloning system, which is recognized as the simplest and most efficient method of cloning PCR products. In this system, single 3′-adenine overhangs are normally first added to the ends of the PCR product so that it becomes complementary to the singly 5′-thymine overhangs on both ends of the expression vector. The complementarity between the linearised vector and its adenylated PCR product then allows for the direct and highly efficiency cloning systems of the two (Zhou and Gomez-Sanchez, 2000). However, the TA cloning system still has its own specific drawback of a possibility of the non-directional cloning system since both ends of the vector are complementary to the either end of the insert (Chen and Janes, 2002).

We had to first verify that the cloned AtCAP-AC gene fragment had been ligated into the pTrcHis2-TOPO expression vector in the correct 5′ to 3′ directional orientation for the proper and successful expression of the correct and desired AtCAP recombinant. Through some specific confirmatory PCR procedures using both primers of the AtCAP-AC insert and then one AtCAP-AC primer plus one pTrcHis2-TOPO vector primer respectively, the amplification of a 360 bp DNA fragment confirmed that the AtCAP-AC gene fragment had indeed been successfully ligated into the pTrcHis2-TOPO while the associated amplification of a 521 bp DNA fragment further confirmed that this ligated AtCAP-AC gene fragment was correctly oriented within the plasmid vector (Figure 3.3). In addition, this correct cloning system of the AtCAP-AC as well as its proper in-frame alignment within the pTrcHis2-TOPO was also further

confirmed and fully validated by the nucleotide-specific sequencing approach (results not shown).

As a result of the various previous studies, it has since been asserted that the cornerstone of modern biotechnology is the expression of functional proteins in heterologous hosts (Gustafsson *et al*., 2004), whereby the heterologous protein expression process itself involves a transfer of the targeted DNA fragments to hosts other than the original source for the subsequent synthesis of such encoded and desired protein products. This in particular is specifically very helpful in the study of plant proteins because protein isolation from plants can, at times, be very cumbersome, costly and lengthy. Thus and in this regard, foreign hosts therefore provide a simpler system for studies on functional and structural properties of specific proteins and an elucidation of their biological and functional roles in complex physiological and cellular processes (Yesilirmak and Sayers, 2009).

Typically, *Escherichia coli* is the most commonly used host for the high-yield expression of most recombinant proteins, usually by exploiting the high promoter specificities and transcriptional activities of its associated bacteriophage T7 RNA polymerase (Robichon *et al*., 2011). Among the variously known *E. coli* host strains, BL21 cells are the most commonly used strain for the recombinant expression and affinity purification of most foreign proteins (Paliy and Gunasekera, 2007). In this regard and in this study therefore, we used the *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cells as our expression host strain of choice because this strain has the same DE3 chromosomal genotype as is the *E. coli* BL21 strain, it also carries a chloramphenicol-resistant pLysS plasmid which encodes for the T7 lysozyme (a natural inhibitor of the T7 RNA polymerase), and it produces a small amount of the T7 lysozyme in order to

suppress the basal expression of the T7 RNA polymerase prior to induction with IPTG (Lucigen Corporation, 2015). Furthermore, as a lysogen of the λDE virus, the *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs' T7 RNA polymerase gene is under the control of an IPTG-inducible *lac*UV5 promoter and the strain is also deficient in the *lon* and *omp*T proteases (Miroux and Walker, 1996).

Based on this, we chose *E. coli* for further recombinant expression and affinity purification of our own targeted and most desired AtCAP-AC protein in this study. Furthermore, the *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs bacterial strain was chosen because of its ability to over-express the recombinant protein upon induction with IPTG, its capacity to prevent the basal expression of the recombinant AtCAP-AC protein as well as for its practical compatibility with the pTrcHis2-TOPO expression vector (which carries the gene fragment downstream of its well-characterised and regulated promoter to produce somewhat relatively high levels of the targeted recombinant protein (Terpe, 2005)).

From this approach and in this study, a subsequent analysis of an SDS-PAGE of the expressed proteins from the non-induced and induced recombinant host cells harbouring the pTrcHis2-TOPO:AtCAP-AC fusion expression construct showed a relatively high level of expression of a protein product of about 22 kDa in the induced cells as compared to the un-induced cells (Figure 3.4). Surprisingly, the correct and expected molecular weight size of the expressed AtCAP-AC recombinant protein product (including the C-terminus 6x-Histidine tag) was 17.8 kDa and not the obtained 22 kDa. In this regard, it is therefore noteworthy that there was actually a typical anomalous gel mobility of this recombinant product (also known as 'gel shifting' or 'migrational shift') that could have been due to one or more of the following possibilities.

Generally, the AtCAP protein is principally annotated as a binding protein, with a molecular function of binding clathrin and/or 1-phosphatidylinositol (Berardini *et al*., 2004) and therefore, by virtue of this natural functional property, it is possible that the expressed recombinant AtCAP-AC could have had bound to one or more other host cell proteins, resulting in the observed anomalous gel shifting. Furthermore, it is also possible that due to this same natural binding property, the recombinant AtCAP-AC could have had somewhat bound to excess SDS molecules than naturally expected during its preparation for SDS-PAGE, and hence resulting in the observed retarded gel mobility (Rath *et al*., 2009). Another possible reason is the existence of many disulphide bridges in this protein, of which such intact bridges have previously been linked to anomalous migrational delays of most proteins because they impose more compact shapes on proteins than usual (Walker and Rapley, 2008).

Furthermore, some amino acid modifications could have had occurred onto the AtCAP-AC recombinant during its preparation for SDS-PAGE, thus resulting in the observed unpredictable electrophoretic mobility on the gel (Shi *et al*., 2012). It is also probable that the β-mercaptoethanol contained in the SDS-PAGE loading buffer could have had exaggerated the effects of heating, leading to the reduced mobility of the recombinant AtCAP-AC in the gel (Sturman, 1977). Lastly, proteins with unusual amino acid sequences, especially those with a high lysine or proline content, very basic proteins, and very acidic proteins, normally behave anomalously in SDS-PAGEs because the charge-to-mass ratios of their SDS-polypeptide complexes are generally different to those that would be expected from their masses alone (AES Application Focus, 2015). Correspondingly, the recombinant AtCAP-AC protein has an overall unusual amino acid composition of about 54%, consisting of a 9.2% lysine content, 4.6% proline content, 16.4% basic amino acid content and a 23.7% acidic amino acid content.

Surprisingly but consistently, when this expressed recombinant AtCAP-AC was later on fully purified from the rest of the other contaminating bacterial proteins (section 5.2.2) followed by its complete refolding and renaturation into a fully functional biological entity (section 5.2.3), its molecular weight size eventually shifted back to the normal and correctly expected level of 17.8 kDa (Figure 5.2) and away from the previously and initially observed 22 kDa, thereby and somewhat consistently validating one or all of the above-given possibilities for its initially and previously observed retarded gel mobility.

Later on and when the expressed recombinant AtCAP-AC protein was assessed for its ability to generate cAMP from ATP, it was found that an induction of its over-expression by 1 mM IPTG resulted in a >3.5-fold increase in the amount of cAMP produced as compared to the non-induced cell cultures (Figure 3.5). This result is somewhat very highly comparable to the one previously reported, whereby when a PSiP coding region of a pollen-specific putative AC from *Zea mays* was cloned into bacterial cells, followed by the treatment of cells with 1 mM IPTG, the treated cells also enhanced their cAMP levels by a factor of >3.0-fold (Moutinho *et al*., 2001). Moreover, the results in Figure 3.5 also show that the presence of forskolin in the induced cells had no apparent effect onto the enzymatic AC activity of the expressed recombinant AtCAP-AC. Thus, while other previous studies have established that the naturally-occurring diterpene, forskolin, could activate transmembrane ACs (tmACs) (Laurenza *et al*., 1989), it has also been shown that all cytosolic soluble ACs (sACs) are totally insensitive to this chemical compound (Forte *et al*., 1983; Yan *et al*., 1998; Buck *et al*., 1999. Thus and as a result of its lack of sensitivity to forskolin, we proposed that the expressed recombinant AtCAP-AC protein could also possibly be a cytosolic sAC.

On the other hand, an exposure of the induced recombinant cells to the P-site effector, dideoxyadenosine, resulted in an overall decrease of the cAMP production of approximately 33% when compared to the recombinant cells induced by IPTG alone (Figure 3.5). These also findings positively correlated with the other previous findings, whereby the 2′,5′-dideoxyadenosine had been shown to inhibit AC activity in isolated membrane preparations, fat cells, broken and intact human platelets, *Xenopus* oocyte membranes, rabbit liver cells, cholera toxin-activated intestinal cells, and the *Zea mays* pollen tube growth protein, at an average level of around 89% when compared to controls (Fain *et al*., 1972; Zenser, 1976; Haslam *et al*., 1978; Londos *et al*., 1978; Florio and Ross, 1983; Sadler and Maller, 1983; Moutinho *et al*., 2001).

3.5 CONCLUSION

The findings of this chapter lead us to propose that the expressed recombinant AtCAP-AC protein is either a *bona fide* cytosolic soluble adenylate cyclase (sAC) capable of generating cAMP from ATP or probably another signaling plant molecule capable of stimulating the resident ACs (*E. coli* ACs in this case) to produce cAMP.

3.6 RECOMMENDATION

Further experimental work on both the *in vitro* and *in vivo* AC activities of the recombinant AtCAP-AC protein would be necessary and pivotal to provide a better and clearer picture as to whether this recombinant protein is indeed a *bona fide* higher plant AC or not.

CHAPTER 4

DETERMINATION OF THE *IN VIVO* ADENYLATE CYCLASE ACTIVITY OF THE RECOMBINANT AtCAP-AC PROTEIN BY COMPLEMENTATION TESTING

ABSTRACT

Despite the significant functional roles played by adenylate cyclases (ACs) in nearly all living organisms, the existence and/or functional involvements of these molecules in higher plants have, for quite some time, remained very elusive. To date, only four higher plant ACs have since been identified, one each in *Zea mays*, *Arabidopsis thaliana*, *Nicotiana benthamiana* and *Hippeastrum hybridum*. However, considering the wide spectrum of processes and activities mediated and/or regulated by the enzymatic product of ACs, adenosine 3′,5′-cyclic monophosphate (cAMP) in plants, it is rather unlikely that all these processes can be mediated by only a single AC across each and every species. Hence, following the recent annotation of a clathrin assembly protein from *Arabidopsis thaliana* (AtCAP) as a possible AC molecule, we hereby describe the recombinant cloning of its AC-containing fragment domain (AtCAP-AC) into a non-lactose fermenting mutant strain of *Escherichia coli* (the *cya*A SP850 cells) and unequivocally confirm its ability to generate cAMP, which in turn then complements the mutant strain to behave as if it was a wild type strain capable of fermenting lactose.

4.1 INTRODUCTION

Adenylate cyclases (ACs) are important molecules involved in the signal transduction pathways of all living organisms as they respond to various environmental stimuli and cues. They convert adenosine 5′-triphosphate (ATP) to cyclic adenosine 3′,5′-monophosphate (cAMP) and a free pyrophosphate group. Cyclic AMP is an important signaling molecule in bacterial species, where it is involved in the regulation of gene expression, and in response to a variety of environmental stimuli and cues (Lory *et al*., 2004).

In *Escherichia coli* specifically, cAMP acts as a second messenger necessary for the synthesis of many inducible enzymes (Pastan and Perlman, 1970), for example, in the operation of its lactose (*lac*) operon (Kuo *et al*., 2003). In this particular system, cAMP forms a complex with the catabolite activator protein and the formed cAMP-CAP complex then binds to the promoter region of the operon. The DNA-bound cAMP-CAP complex in turn, physically interacts with RNA polymerase and increases its affinity for the *lac* promoter, hence activating the transcriptional process of the *lac* operon (de Gunzburg, 1985). In the presence of lactose, the lactose repressor generally falls off the *lac* operator site, thus allowing RNA polymerase to bind to the operator, leading to the transcription of the β-galactosidase, lactose permease and transacetylase genes of the operon. This process is normally aided by a background level of the β-galactosidase which basally converts some lactose to allolactose, the inducer of the *lac* operon (Hames and Hooper, 2005). Cyclic AMP and ACs in this regard, therefore play a central role in the physiological processes of *E. coli*.

In this bacterial strain (*E. coli*), only a single AC naturally exists and is encoded for by the *cya*A gene (Côté *et al*., 2010), and this one and only enzyme was first isolated and purified by Yang

and Epstein in 1983. Later on, it was then found that the *Bordetella pertussis cya*A does encode for an AC toxin (Gross *et al*., 1992). In 1992, Gross *et al.* then used a site-directed (deletion) mutagenesis and genetic recombination process that specifically was targeted at the *B. pertussis cya*A gene and in order to produce mutants that did not have the AC activity of the toxin. Subsequently, it was then reported that an *E. coli cya*A mutant strain, known as the SP850, had a *cya*A mutation and threfore lacked the AC activity (Cotta *et al*., 1998).

This mutant strain can not ferment carbon sources such as lactose, maltose, galactose, arabinose and glycerol, and it also grows very slowly on glucose, fructose and galactose sugars (Perlman and Pastan, 1969). The mutant is actually deficient in the *lacZ* expression because it lacks the AC needed for the cAMP production and therefore, cannot activate the catabolite activator protein, meaning that its cAMP receptor protein-dependent promoter is presumably inactive (Dole *et al*., 2004). The mutant also possesses some resistance to kanamycin (Cotta *et al*., 1998), is pH-sensitive for growth (Ahmad and Newman, 1988), and is a vibrant host for the recombinant expression systems of foreign proteins, especially when it is in its exponential phase (Grossman *et al*., 1998).

Technically, the *cyaA* SP850 mutant provides a viable opportunity for testing the possible AC activities of foreign recombinant proteins via a complementation phenomenon of the mutation. In genetic analysis, a complementation test is used in order to determine if two recessive mutations that produce the same phenotype are located on the same gene (Hanoune and Defer, 2001). Thus through this same test, the *in vivo* AC activity of unknown recombinant protein molecules can then be pre-screened and the promising candidates subsequently further evaluated via some specific *in vitro* immunoassaying systems (Ruzvidzo *et al*., 2013).

While a wild type *E. coli* strain naturally produces a deep red colour on MacConkey lactose agar and as a result of its ability to ferment lactose, the SP850 strain, due to its *cya*A mutation, cannot utilize lactose and therefore produces colourless colonies on the same growth medium. Apparently, when this mutant strain is transformed with any foreign gene capable of encoding for a functional AC activity, it also produces a deep red colour on the MacConkey lactose agar, which is indicative of a complementation of its AC deletion (Cotta *et al*., 1998), leading to the production of the most needed cAMP, which in turn, then restores the normal utilization of lactose as a carbon source (Perlman and Pastan, 1969).

Based on this phenomenon, 2001, Moutinho *et al*. conducted a complementation process of the *cya*A mutation in the *E. coli* SP850 strain with the PSiP protein from *Zea mays* and, in this study, the presence of a functional recombinant protein successfully gave rise to a selectable phenotype of the deep purple colonies. This specific finding was also further confirmed and fully validated through a similar approach by Ruzvidzo *et al*. in 2013 and Swiezawska *et al*. in 2014 using the AtPPR-AC and HpAC1 proteins respectively. In the same vein, Charania *et al*. (2009) also similarly demonstrated that the *Shewanella oneidensis cya*A, *cya*B and *cya*C genes predicted to encode for the classes I, III and IV, do indeed and actually encode for proteins with functional AC activities in the SP850 *cya*A mutant.

Thus following a recent annotation of the Arabidopsis clathrin assembly protein (AtCAP), encoded for by the At1g68110 gene, as a putative AC molecule (Gehring, 2010), this present study therefore, attempted to provide some practical evidence of the *in vivo* AC activities of this putative protein through a complementation process of the SP850 *E. coli* mutant strain by its AC-harbouring fragment domain, the AtCAP-AC.

4.2 METHODOLOGY

The ability of the AtCAP-AC recombinant protein to produce cAMP from ATP *in vivo* was determined via a complementation process of the mutant *E. coli* host strain (the *cya*A SP850) (Coli Genetic Stock Center, Yale University, Connecticut, USA) with the pTrcHis2-TOPO:AtCAP-AC fusion construct, followed by a visual screening for a selectable phenotype.

4.2.1 ISOLATION AND PURIFICATION OF THE pTrcHis2-TOPO:AtCAP-AC EXPRESSION CONSTRUCT FROM THE *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs CELLS

The pTrcHis2-TOPO:AtCAP-AC fusion expression construct that was previously confirmed to be correctly and properly designed in section 3.2.5.4.2 and then later on maintained in the chemically competent *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs expression cells (section 3.2.5.5) was extracted as previously outlined in section 3.2.5.4.1 using a GeneJET Plasmid Miniprep Kit and in accordance with the manufacturer's protocol and instructions (Catalogue # K0503; ThermoScientific Inc., New York, USA). The resultant pTrcHis2-TOPO:AtCAP-AC fusion expression construct was then later on used to transform the chemically competent mutant SP850 host cells for the undertaking of the intended complementation testing process.

4.2.2 PREPARATION OF COMPETENT COMPLEMENTATION HOST CELLS AND THEIR TRANSFORMATION WITH THE pTrcHis2-TOPO:AtCAP-AC EXPRESSION CONSTRUCT

The *E. coli cya*A SP850 mutant host cells (Coli Genetic Stock Center, Yale University, Connecticut, USA) were first prepared to become chemically competent before being transformed with the pTrcHis2-TOPO:AtCAP-AC fusion expression construct.

4.2.2.1 Induction of the Chemical Competence into the *E. coli cya*A SP850 Cells

A 10 ml overnight culture of the *E. coli cya*A SP850 cells was prepared by inoculating a selective LB broth (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) sodium chloride, 0.5% (w/v) glucose, 15 µg/ml kanamycin, pH: 7.0) with the SP850 cells and in order to enrich them to their stationary phase. The next morning, 20 ml of the fresh pre-warmed LB broth supplemented with 15 µg/ml kanamycin was inoculated with 1 ml of the overnight culture and incubated at 37ºC, at 200 rpm, and up until an OD_{600} of 0.5 was been reached. Cells were then harvested by centrifuging at 4,000*g* for 5 minutes at 4ºC. The supernatant was discarded while the harvested cells were resuspended in 15 ml of an ice-cold Transformation Buffer 1 (30 mM potassium acetate, 50 mM manganese chloride, 100 mM rubidium chloride, 10 mM calcium chloride, 15% (v/v) sterile glycerol, pH: 5.8) and incubated on ice for 90 minutes. The cells were then harvested by centrifuging at 4,000*g* for 5 minutes at 4ºC. The supernatant was discarded while the harvested cells were resuspended in 4 ml of an ice-cold Transformation Buffer 2 (10 mM MOPS [3-(N-morpholino) propanesulfonic acid], 75 mM calcium chloride, 10 mM rubidium chloride, 15% (v/v) glycerol, pH: 6.8). The resuspended cells were then aliquoted into 100 µl portions and kept on ice before being used in the subsequent step.

4.2.2.2 Transformation of the Chemically Competent *E. coli cya*A SP850 Cells with the pTrcHis2-TOPO:AtCAP-AC Fusion Expression Construct

About 10 µl of the pTrcHis2-TOPO:AtCAP-AC fusion expression construct were aseptically transferred into an ice-cold 1.5 ml microfuge tube containing the 100 µl portion of the chemically competent SP850 *E. coli* mutant cells and the mixture was gently mixed before being incubated for 20 minutes on ice. The mixture was then heat-shocked through incubation at 42ºC

for 90 seconds on a dry bath heating block before being incubated on ice for 5 minutes. The transformation mixture was then supplemented with 500 µl of the SOC medium and incubated for 90 minutes at 37ºC in an orbital shaker at 200 rpm. Using the spread-plate method, 100 µl aliquots of the transformation mixture were then inoculated onto LB agar plates supplemented with 15 µg/ml kanamycin and 100 µg/ml ampicillin. The innoculated plates were then incubated overnight at 37ºC and the generated cell colonies used in the next step.

4.2.3 Cell Culturing and Phenotypic Scoring

A selective MacConkey lactose agar plate (2% (w/v) peptone, 1% (w/v) lactose, 0.5% (w/v) bile salts, 0.0075% (w/v) neutral red, 1.5% (w/v) agar, 15 µg/ml kanamycin and 0.1 mM IPTG) was divided into three quadrants (Figure 4.1). Using the streak-plate method, quadrant 3 was streaked with the *cya*A cells transformed with the pTrcHis2-TOPO:AtCAP-AC fusion expression construct, quadrant 2 with the non-transformed *cya*A mutant cells, while quadrant 1 was left unstreaked (Figure 4.1). The plate was then incubated for 48 hours at 37ºC, and its quadrants later on analysed for the visually different and scorable phenotypic characteristics.

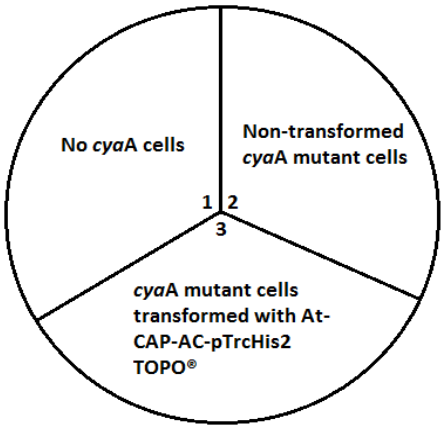

Figure 4.1: Culture distribution of the *cya*A SP850 mutant cells on a MacConkey agar plate. Segment 1 contains no bacterial cells, segment 2 contains the non-transformed *cya*A mutant cells, while segment 3 contains the *cya*A mutant cells harbouring the AtCAP-AC gene fragment.

4.3 RESULTS

In order to determine the *in vivo* adenylate cyclase activity of the cloned recombinant AtCAP-AC gene fragment, a complementation test was conducted using the *E. coli cya*A SP850 mutant host strain (Coli Genetic Stock Center, Yale University, Connecticut, USA) which systematically lacks the inherent AC activity. In this test, the mutant cells were transformed with the pTrcHis2-TOPO fusion expression vector harbouring the AtCAP-AC gene fragment (pTrcHis2-TOPO:AtCAP-AC) followed by an analysis of the AtCAP-AC's ability to generate cAMP on a differential and selective MacConkey agar plate, and as is shown in Figure 4.2 below.

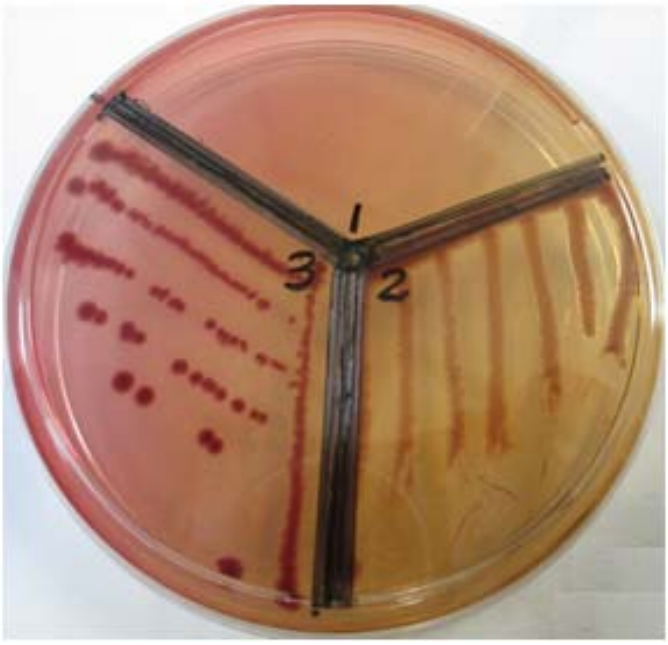

Figure 4.2: Determination of the *in vivo* enzymatic AC activity of the recombinant AtCAP-AC protein via a complementation test. Segment 1 of the plate contains no bacterial cells, segment 2 contains the non-transformed *E. coli cya*A SP850 mutant cells while segment 3 contains the mutant cells harbouring the recombinant pTrcHis2-TOPO:AtCAP-AC fusion expression construct. The non-transformed cells in segment 2 are non-lactose fermenters and therefore, produced whitish/yellowish colonies, while on the other hand, the recombinant cells in segment 3 exhibited a magenta/deep purple phenotype that signifies a lactose-fermenting phenotype.

4.4 DISCUSSION

The *Escherichia coli cya*A SP850 mutant strain was created by transducing a total gene deletion of the *cya* gene from the SP793 strain into the CA8000 strain via a P1 background transduction system (Shah and Peterkofsky, 1991). In this experiment, it was shown that the newly generated SP850 strain did exhibit all the properties expected for the lack of AC activity and therefore, was subsequently recommended for use as a clean *cya* deletion, before its eventual deposition into the *E. coli* Genetic Stock Center of the Yale University. This thus therefore made the SP850 strain a perfect host strain for all subsequent complementation assays seeking to test the *in vivo* AC activities of any unknown recombinant and putative AC protein candidate.

In this study, we conducted a complementation test of the AC fragment domain of a clathrin assembly protein from *A. thaliana* (AtCAP-AC) that previously had been bioinformatically annotated as a putative AC (Gehring, 2010). We constructed its expression system in the form of a pTrcHis2-TOPO:AtCAP-AC fusion construct and used this expression system to transform the chemically competent *E. coli cya*A SP850 mutant host cells followed by a visual analysis of the scorable cell phenotypes on a selective and differential MacConkey agar medium. From this set up, we found out that the cells harbouring the pTrcHis2-TOPO:AtCAP-AC fusion construct appeared deep red/purple as opposed to those that were not transformed, which appeared white coloured/pale yellow. This outcome therefore conclusively affirmed the recombinant AtCAP-AC protein as a *bona fide* higher plant AC, capable of generating cAMP from ATP, which in turn was responsible for the fermentation of the lactose sugar and resulting in the production of the observed deep red/purple cell colonies.

Essentially and in all wild type *E. coli* cells capable of generating cAMP, when the cAMP is bound to the transcriptional catabolite gene-activator protein (CGAP), it acts as a pleiotropic regulator for the expression of genes involved in the catabolism of carbohydrates such as lactose and maltose. Hence, any *E. coli* strain lacking cAMP (mutant) is systematically unable to ferment these sugars, a condition that can practically be reversed by expressing some putative AC protein candidates within this mutant (Moutinho *et al*., 2001), and as was successfully done in this study.

To date, the practical complementation of the *E. coli cya*A SP850 mutant strain has so far been successfully undertaken with seven putative AC protein candidates, four from the anaerobic bacteria and three from higher plants. Firstly, a recombinant cloning of the *Prevotella ruminicola* D31d protein (the first AC gene to be isolated from an anaerobic bacterium) into the *E. coli cya*A SP850 mutant cells resulted in the growth of some deep red cell colonies on the MacConkey lactose agar, which was indicative of lactose fermentation and a complementation of the AC deletion in the mutant cells (Cotta *et al*., 1998). Secondly, Moutinho *et al*. (2001) demonstrated that the pollen signaling protein (PSiP) from *Zea mays* possessed an AC activity by cloning it into the *E. coli cya*A SP850 mutant cells, which in turn, resulted in the growth of some deep purple-stained cells on MacConkey lactose agar. Thirdly, Charania *et al*. (2009) functionally validated the AC activities of the *cyaA, cyaB* and *cyaC* genes from the anaerobic bacterium *Shewanella oneidensis* MR-1 through a complementation of the SP850 mutant. Fourthly, Ruzvidzo *et al*. (2013) cloned the AC fragment domain of a pentatricopeptide protein from *Arabidopsis thaliana* (AtPPR-AC) into the SP850 mutant and reported the exhibition of a magenta coloured phenotype of the recombinant cells on MacConkey lactose agar. Lastly and

more recently, another plant AC, the HpAC1 from *Hippeastrum hybridum*, explicitly exhibited an ability to compensate for the AC deficiency in the SP850 mutant (Swiezawska *et al*., 2014).

Conceivably, since our preceding work in this study on the endogenous activity assaying of the recombinant AtCAP-AC protein (Chapter 3) had failed to conclusively establish this putattive protein as a *bona fide* AC molecule, findings of this chapter now therefore firmly provide strong and undebatable evidence that this recombinant protein is indeed a *bona fide* functional AC molecule capable of generating cAMP from ATP, and possibly involved in cell-cell communication signalling systems that are mediated and/or regulated by cAMP.

4.5 CONCLUSION

The findings of our complementation testing in this study unequivocally and firmly established the Arabidopsis clathrin assembly protein (AtCAP) as a *bona fide* and biologically functional AC molecule and thereby becoming the fifth ever such candidate to be identified and experimentally confirmed in higher plants, while it also becomes the second ever such molecule to be identified in the *Arabidopsis thaliana*.

4.6 RECOMMENDATION

Despite the establishment of the Arabidopsis clathrin assembly protein (AtCAP) as a *bona fide* higher plant AC molecule in this chapter, its molecular properties and possible exact physiological roles in plants have not yet been fully elucidated and therefore, further experimental work on the recombinant AtCAP-AC is still necessary to attempt and establish all this missing information.

CHAPTER 5

AFFINITY PURIFICATION AND FUNCTIONAL CHARACTERISATION OF THE RECOMBINANT AtCAP-AC PROTEIN

ABSTRACT

For several years, the demonstration of potential adenylate cyclase (AC) activities in higher plants has always been conducted with the histochemical and/or biochemical methods. However, these traditional methods have since been seriously criticized on the basis of their strong reliance on either intact cells and/or crude cellular extracts, which in turn always resulted in the final findings being more of rather presumptive deductions than firm or conclusive outcomes. It thus has become very apparent that it is necessary to use pure protein extracts so as to provide a more convincing and highly reliable functional evidence for this very important group of biological molecules in plants. Following the recent annotation of a clathrin assembly protein (AtCAP) encoded by the At1g68110 gene from *Arabidopsis thaliana* as a possible AC candidate, we hereby detail the recombinant expression and affinity purification of its AC-containing fragment domain (AtCAP-AC) followed by a practical demonstration of its *in vitro* enzymatic activity. In addition, we also show that this truncated recombinant protein is indeed a soluble AC (sAC), whose functional activities may be mediated by cAMP and possibly, via a calmodulin dependent signalling system.

5.1 INTRODUCTION

The natural occurrence of cyclic nucleotides and nucleotide cyclases in plants was once strongly doubted because the levels of both cAMP and cGMP in higher plants are very low and therefore, it has been a very huge challenge particularly, to demonstrate the presence of nucleotide cyclases in plants (Schaap, 2005). Apparently, even though the cytosolic secondary messengers (cAMP and cGMP) are known to be involved in numerous signal transduction pathways in plants (Leng *et al*., 1999), the presence and/or existence of the nucleotide cyclases that specifically generate these secondary messengers have still been very elusive. The new frontier in this respect is therefore, now to characterise and elucidate the functional roles of nucleotide cyclases (guanylate cyclases (GCs) and adenylate cyclases (ACs)) in plant signal transduction systems.

To date, a number of studies have been conducted which successfully identified and characterised GCs in plants. The AtGC1 protein from *Arabidopsis thaliana* has been shown to possess a magnesium-dependent *in vitro* GC activity which is not nitrous oxide-dependent, and to also have a slight AC activity that was partly dependent on both magnesium and manganese as co-factors of activity (Ludidi and Gehring, 2002). The AtBRI1 protein from *Arabidopsis thaliana* has also been shown to possess an *in vitro* GC activity (Kwezi *et al*., 2007). Subsequently, the characterisation of an AtPepR1 from Arabidopsis also showed that its cytoplasmic domain could acts as a GC *in vitro* that prefers Mn^{2+} to Mg^{2+} as a co-factor, and that it activates a plasma membrane-localised cyclic nucleotide-gated calcium-conducting channel as well as stimulating the cyclic nucleotide-gated channel-dependent cytosolic calcium elevations (Qi *et al*., 2010). In addition, the PSKR1 protein from Arabidopsis was also identified as a functional GC which enables a cGMP-dependent signalling (Kwezi *et al*., 2011). Eventually,

upon the characterisation of a PSKR1, Kwezi *et al*. (2011) also found that the protein is specific for cGMP production and had no specific preference for either magnesium or manganese as the crucial co-factor of activity. Furthermore, studies on the AtNOGC1 from *Arabidopsis thaliana* showed that it has *in vitro* GC activity, has preference for Mn^{2+} to Mg^{2+} as a co-factor, has a substrate specificity for GTP and not ATP, binds to nitric oxide with a higher affinity than it binds to oxygen, is nitrous oxide-dependent, and its expression is elevated during the nitric oxide starvations (Mulaudzi *et al*., 2011).

With regard to plant ACs, Gehring (2010) bioinformatically identified nine putative AC candidates in the Arabidopsis genome. To date however, only four ACs have yet been experimentally tested and functionally confirmed in different plant species. Firstly, Moutinho *et al*. (2001) reported that a *Zea mays* pollen signaling protein (PSiP) possesses an *in vivo* AC activity and is responsible for the polarized growth of pollen tubes. Secondly, a pentatricopeptide protein from *Arabidopsis thaliana* (AtPPR; At1g62590) has also been shown to possess both the *in vitro* and *in vivo* AC activities (Ruzvidzo *et al*., 2013). Thirdly, studies on a protein from *Nicotiana benthamiana*, NbAC, showed that the protein has an *in vitro* AC activity that participates in the tabtoxinine-β-lactam-induced cell death which occurs during the development of the necrotic wildfire disease (Ito *et al*., 2014). Lastly, the HpAC1 protein from *Hippeastrum hybridum* has been shown to possess an *in vivo* AC activity and that its expression is significantly increased in response to mechanical damage and the *Phoma narcissi* infection (Swiezawska *et al*., 2014). Thus, out of the total of four ACs that so far have been experimentally confirmed, three have already been functionally characterised (Moutinho *et al*., 2001; Ito *et al*., 2014; Swiezawska *et al*., 2014).

Meanwhile, of the nine Arabidopsis proteins that previously had been bioinformatically annotated as putative ACs, the clathrin assembly protein (AtCAP) encoded by the At1g68110 gene has since been implicated in the signal transduction process of stress response (De Vos *et al*., 2005; Schmid *et al*., 2005; De Vos and Jander, 2009) via cAMP (Gehring, 2010). Apparently, while we have succeeded to provide full experimental evidence for the endogenous and *in vivo* AC activities of the AtCAP-AC protein in Chapters 3 and 4 above respectively, there is still an apparent need to further determine how these activities are mediated and/or modulated at the molecular level. In this chapter therefore, we detail the steps taken in purifying, refolding and biochemically characterizing the recombinant AtCAP-AC protein.

5.2 METHODOLOGY

5.2.1 DETERMINATION OF THE SOLUBLE/INSOLUBLE NATURE OF THE RECOMBINANT AtCAP-AC PROTEIN

The solubility/insolubility nature of the recombinant AtCAP-AC protein was determined under native non-denaturing conditions by first lysing the induced recombinant cells followed by an analysis of their cellular protein components through SDS-PAGE.

5.2.1.1 Over-expression of the Recombinant AtCAP-AC Protein

In order to yield high and enhanced levels of the recombinant AtCAP-AC protein, its expression profile in section 3.2.6.1 was strategically modified, whereby the cell culture at an OD_{600} of 0.5 was this time around split into two sets (the control culture and the induced culture) instead of four, and the induced culture then allowed to grow for a further 3 hours instead of 1, at 37ºC at 200 rpm.

5.2.1.2 Determination of the Soluble/Insoluble Nature of the Recombinant AtCAP-AC Protein

A total volume of 10 ml of the induced *E. cloni* EXPRESS BL21 (DE3) pLysS DUOs cells was centrifuged at 9,200*g* for 5 minutes. The supernatant was discarded while the pelleted cells were resuspended in 2.5 ml of sterile Tris-buffered saline (TBS) solution (50 mM Tris-HCl, 150 mM NaCl, pH: 7.5) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 µg/ml lysozyme (Catalogue # 62970; Sigma-Aldrich Inc., Missouri, USA). The mixture was incubated on ice for 1 hour and then periodically mixed by vortexing at medium speed for 10 minutes with 1 minute intervals of ice incubation. The mixture was then centrifuged at 9,200*g* for 2 minutes and both the supernatant and its pellet were then separately kept as the soluble and insoluble protein fractions respectively. An aliquot (20 µl) of each of these two fractions was then analyzed by SDS-PAGE and in order to determine the presence and relative amount of the expressed recombinant AtCAP-AC in each one of them. A presence of the recombinant protein in the supernatant fraction would indicate its soluble nature while its presence in the pellet fraction would, on the other hand, indicate its insoluble nature.

5.2.2 PURIFICATION OF THE RECOMBINANT AtCAP-AC PROTEIN

After realizing that the targeted recombinant AtCAP-AC protein has been wholly expressed as an insoluble product, we then therefore purified it under non-native denaturing conditions using a HIS-Select nickel-nitrilotriacetic acid (Ni-NTA) affinity matrix and in accordance with the manufacturer's protocol (Catalogue # P6611; Sigma-Aldrich Inc., Missouri, USA) as is briefly described below.

5.2.2.1 Preparation of the Cleared Lysate

The cell pellet component (insoluble fraction) generated in section 5.2.1.2 was resuspended in 5 ml of Lysis Buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris-HCl; pH: 8.0, 500 mM NaCl, 20 mM β-mercaptoethanol, 10 mM imidazole and 7.5% (v/v) glycerol) at a ratio of 5 ml buffer:1 g pellet weight. The cell pellet component was then solubilized into a crude lysate through vigorous vortexing at full speed for 60 minutes. The crude lysate was then centrifuged at 10,000g for 30 minutes at room temperature to yield a cleared lysate (supernatant) through its separation from the overall cell debris. The yielded cleared lysate was then collected and stored for further downstream processing.

5.2.2.2 Equilibration of the HIS-Select Ni-NTA Affinity Matrix

A 5 ml aliquot of the 50% (v/v) HIS-Select Ni-NTA affinity bead solution was allowed to settle down at room temperature in a sterile 15 ml falcon tube before 2.5 ml of the 30% (v/v) storage ethanol was completely removed. The beads were then washed twice with 2.5 ml of sterile distilled water and in order to remove all the residual storage ethanol. The beads were then equilibrated with 5 ml of Equilibration Buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris-HCl; pH: 8.0, 500 mM NaCl, 20 mM β-mercaptoethanol, 7.5% (v/v) glycerol) by mixing the two and allowing the mixture to settle down at room temperature for 5 minutes before the buffer was completely removed and discarded.

5.2.2.3 Binding of the Recombinant AtCAP-AC Protein onto the HIS-Select Ni-NTA Affinity Matrix

The 5 ml cleared lysate produced in section 5.2.2.1 was transferred to the 2.5 ml beads equilibrated in section 5.2.2.2 and the mixture was then allowed to gently mix for 1 hour at room temperature on an Adjustable Rotator Revolver (Labnet International Inc., New Jersey, USA) at 30 rpm. During this phase, the recombinant AtCAP-AC protein then became bound to the HIS-Select Ni-NTA affinity beads. The mixture was then centrifuged at 5,000*g* for 5 minutes before the supernatant (flow-through) was removed. About 20 µl of the flow-through was then saved for analysis by SDS-PAGE while the rest of it was discarded.

5.2.2.4 Washing of the Bound HIS-Select Ni-NTA Affinity Matrix

After binding the recombinant AtCAP-AC protein onto the equilibrated HIS-Select Ni-NTA affinity matrix, the beads were washed thrice to remove all the unbound bacterial proteins. Two column volumes (cv or 5 ml) of Wash Buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris-HCl; pH: 8.0, 500 mM NaCl, 20 mM β-mercaptoethanol, 7.5% (v/v) glycerol, and 10 mM imidazole) were added to the bound affinity beads and the suspension was mixed on a bench revolver at 30 rpm for 5 minutes. The suspension was then centrifuged for 30 seconds at 5,000*g* before the supernatant (wash 1) was removed and an aliquot saved for analysis by SDS-PAGE. The washing step was repeated two more times (wash 2 and 3) before a 20 µl aliquot of the washed bound beads was also collected and saved for analysis by SDS-PAGE. Additionally, both washes 2 and 3 (20 µl aliquots) were also saved for analysis by SDS-PAGE.

5.2.3 REFOLDING OF THE PURIFIED AND DENATURED RECOMBINANT AtCAP-AC PROTEIN

In order to convert the purified recombinant AtCAP-AC protein from its denatured and non-functional state into its native and functional form, the following refolding procedures were sequentially undertaken:

5.2.3.1 Preparation of the Refolding Column

The washed HIS-Select Ni-NTA affinity beads carrying the bound and purified recombinant AtCAP-AC protein were first resuspended in 10 cv of Gradient Buffer (8 M urea, 200 mM NaCl, 50 mM Tris-HCl; pH: 8.0, and 20 mM β-mercaptoethanol) and the suspension slurry was then loaded into an empty XK16 column (Bio-Rad Laboratories Inc., California, USA) with the column tap at the bottom closed. The beads were then allowed to sediment by draining out 8 cv of the buffer through the tap before the loaded column was connected to a BioLogic DuoFlow Chromatography System (Bio-Rad Laboratories Inc., California, USA) for the initiation of the gradient refolding process.

5.2.3.2 The Refolding Gradient System

A linearized refolding gradient system was created and run on the BioLogic DuoFlow Chromatography System based on the parameters listed in Table 5.1 below. In this system, the 8 M urea gradient buffer was slowly and linearly diluted to 0 M urea concentration with a Refolding Buffer (200 mM NaCl, 50 mM Tris-HCl; pH: 8.0, 500 mM glucose, 0.05% (w/v) polyethylglycol, 4 mM reduced glutathione, 0.04 mM oxidized glutathione, 100 mM non-detergent sulfobateine, and 0.5 mM PMSF). Ideally, this process subsequently converted the

purified AtCAP-AC recombinant protein from its denatured and non-functional state into a native and biologically active molecule.

Table 5.1: Conditions for the refolding process of the recombinant AtCAP-AC protein using the BioLogic DuoFlow Chromatography System.

Variable	Value
Column volume	1.00 ml
Flow rate	0.50 ml/min
Column pressure limit	2.80 MPa
Average time for UV	1.00 sec
System pump for automatic pressure and flow regulation	Normal
Starting concentration of buffer B (Refolding Buffer)	0.00%
Target concentration of buffer A (Lysis Buffer)	100.00%
Equilibrate column with	2.00 cv
Length of gradient	600 min
Clean after gradient with (Washing Buffer)	10.00 ml
Elute protein with (Elution Buffer)	1.00 ml

5.2.3.3 Elution of the Refolded Recombinant AtCAP-AC Protein

In order to elute the refolded recombinant AtCAP-AC protein from the HIS-Select Ni-NTA affinity matrix, the bound beads in the column were first washed with 10 cv of Washing Buffer (200 mM NaCl, 50 mM Tris-HCl; pH: 8.0, 20% (v/v) glycerol, 0.5 mM PMSF), and also in order to remove any residual refolding agents from the refolded protein. The washed and refolded recombinant AtCAP-AC protein was then eluted in 2 ml Elution Buffer (200 mM NaCl, 50 mM Tris-HCl; pH 8.0, 250 mM imidazole, 0.5 mM PMSF, and 20% (v/v) glycerol), using an elution process that was directly linked to the DuoFlow system.

5.2.3.4 Concentration and De-salting of the Recombinant AtCAP-AC Protein

The eluted recombinant AtCAP-AC protein sample (2 ml) was both concentrated and de-salted by transferring it into a Spin-XUF concentrating and de-salting device with a molecular weight cut off point of 3.0 kDa (Corning Corp., New York, USA), followed by centrifugation at 2,540*g* (Hermle Labortechnik, Wehingen, Germany) at 4ºC and up until the final volume had gone down to 200 µl. The concentrated and de-salted protein fraction was then transferred into a new sterile 1.5 ml microfuge tube and protein concentration determined with a 2000 Nanodrop Spectrophotometer (Thermo Scientific Inc., Massachusettes, USA). A small fraction (20 µl) of the eluted protein fraction was then set aside for analysis by SDS-PAGE while the rest was stored at -20ºC for further downstream use.

5.2.4 FUNCTIONAL CHARACTERISATION OF THE RECOMBINANT AtCAP-AC PROTEIN

An *in vitro* assay to determine and characterise the AC activity of the purified recombinant AtCAP-AC protein was conducted via a cAMP-linked enzyme immunoassaying system that was based on its acetylation protocol and in accordance with the manufacturer's instructions (Catalogue # CA201; Sigma-Aldrich Inc., Missouri, USA). The assaying process included a determination of the functional effects of different ions and various molecules (ATP, GTP, Mg^{2+}, Mn^{2+}, Ca^{2+}, HCO_3^-, and F^-) on the enzymatic activity of the purified recombinant protein. The *in vitro* assay was conducted in triplicate form and the results subjected to an analysis of variance (Super-Anova, Stats Graphics Version 7, 1993, Stats Graphics Corporation, USA).

5.2.4.1 Preparation of Samples and Enzyme Immunoassaying

Several 200 µl reaction mixes were prepared in sterile 1.5 ml microfuge tubes and as is detailed in Table 5.2 below. Each tube was made to contain 50 mM of Tris-HCl (pH: 8.0) for buffering purposes as well as 2 mM of IBMX to prevent the degradation of cAMP by phosphodiesterases. Each reaction mix was then incubated at room temperature for 20 minutes and the reaction system stopped by adding 1 mM EDTA to chelate out all the divalent metal ions and hence removing the co-factors necessary for the enzymatic action of the recombinant AtCAP-AC. Additionally, the samples were boiled for 5 minutes in order to inactivate the protein. The inactivated samples were then clarified by centrifugation at 16,300*g* for 5 minutes. Their cAMP contents were then acetylated by adding the acetylating reagent (2:1, triethylamine:acetic anhydride (v/v)) at a volumetric ratio of 1:20 acetylating reagent:sample. The acetylation reaction solutions were then vigorously mixed through vortexing at high speed for 2 seconds and their inherent cAMP levels subsequently measured by enzyme immunoassaying.

Table 5.2: Molecular characterisation of the recombinant AtCAP-AC protein.

	Tube 1	**Tube 2**	**Tube 3**	**Tube 4**	**Tube 5**	**Tube 6**	**Tube 7**	**Tube 8**	**Tube 9**
Tris	50 Mm	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM	50mM
IBMX	2 mM	2 mM	2 mM	2 mM	2 Mm	2 mM	2 mM	2 mM	2mM
$MgCl_2$	5 mM	5 mM	-	5 mM	5 mM	5 mM	5 mM	5 mM	5mM
Protein	-	25 µg	25 µg	25 µg	25 µg	25 µg	25 µg	25 µg	25µg
ATP	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	-	1 mM	1mM
$MnCl_2$	-	-	5 mM	-	-	-	-	-	-
$CaCl_2$	-	-	-	250 µM	-	-	-	-	-
GTP	-	-	-	-	-	1mM	1mM	-	-
$NaHCO_3$	-	-	-	-	50 mM	-	-	-	-
NaF + $AlCl_3$	-	-	-	-	-	-	-	-	10 mM + 30 µM

5.2.5 STATISTICAL ANALYSIS OF ENZYME IMMUNOASSAYING DATA

The results of the enzyme immunoassaying system were all based on the means of three replicates whereby all outcomes of each assay were subjected to analysis of variance (ANOVA) (Super-Anova, Statsgraphics Version 7, 1993, Statsgraphics Corporation, USA). Wherever the ANOVA revealed some significant differences between outcomes, then the means (n = 3) were separated with a *post hoc* Student Newman-Kuehls (SNK), multiple range test ($p \leq 0.05$).

5.3 RESULTS

5.3.1 Purification of the Recombinant AtCAP-AC Protein

Expression of recombinant proteins in various prokaryotic systems such as *E.coli* usually results in aggregation of the recombinant protein into inclusion bodies (Rudolph and Lilie, 1996). In this study and after determining that the recombinant AtCAP-AC protein was wholly expressed in form of inclusion bodies, this poly-histidine C-tagged fusion product was then purified under non-native denaturing conditions using a HIS-Select Ni-NTA affinity matrix system (Hochuli *et al.*, 1987) and based on the manufacturer's protocol and instructions (Catalogue # P6611; Sigma-Aldrich Inc., Missouri, USA) as is shown in Figure 5.1 below.

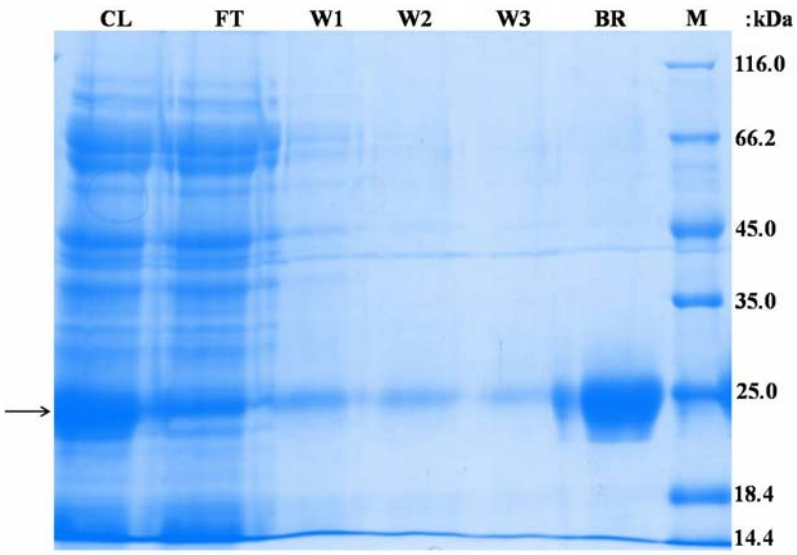

Figure 5.1: Purification of the recombinant AtCAP-AC protein under non-native denaturing conditions. An SDS-PAGE resolution of various fractions of the recombinant AtCAP-AC protein that were collected at the different stages of its purification process. **CL** represents the cleared lysate generated from the inclusion bodies after their solubilisation in 8 M urea buffer; **FT** represents the flow-through of the cleared lysate after had passed through the equilibrated HIS-Select Ni-NTA matrix; **W1**, **W2** and **W3** represent the three successive washes of the bound AtCAP-AC protein onto the matrix; and **BR** represents the purified but bound AtCAP-AC protein. The arrow marks the recombinant AtCAP-AC protein product while **M** represents the unstained low molecular weight marker (ThermoFisher Scientific Inc., New York, USA).

5.3.2 Refolding and Elution of the Recombinant AtCAP-AC Protein

Following the purification process of the recombinant AtCAP-AC protein under non-native denaturing conditions, it was necessary for the denatured and non-functional recombinant protein to be facilitated to regain its biological activity. While it was still bound onto the HIS-Select Ni-NTA affinity matrix, the purified recombinant AtCAP-AC was refolded to its native and functional form via a linear refolding gradient system, and using the DuoFlow BioLogic system (Bio-Rad Laboratories Inc., California, USA). In this system, the 8 M urea buffer into which the recombinant AtCAP-AC was initially suspended was diluted over a 600-minute time period with

a non-denaturing buffer to a final concentration of 0 M urea and thereby refolding the protein back into its active state. Subsequently, the expected 17.8 kDa refolded and purified recombinant AtCAP-AC protein was eluted from the HIS-Select Ni-NTA affinity matrix (Figure 3.2), de-salted and concentrated in preparation for the enzyme immunoassaying system.

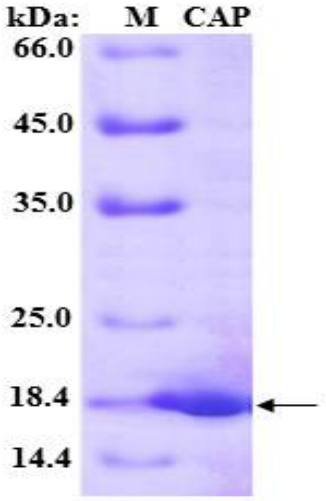

Figure 5.2: Refolding and elution of the purified recombinant AtCAP-AC protein. An SDS-PAGE resolution of the refolded and purified recombinant AtCAP-AC protein eluted from the HIS-Select Ni-NTA affinity matrix, where M represents the low molecular weight marker (ThermoFisher Scientific Inc., Missouri, USA) and the arrow marking the refolded, purified and eluted recombinant AtCAP-AC protein product.

5.3.3 Functional Characterisation of the Recombinant AtCAP-AC Protein

An *in vitro* assaying system of the recombinant AtCAP-AC protein using the cAMP-linked Enzyme Immunoassay System (Catalog # CA201; Sigma-Aldrich Inc., Missouri, USA) was conducted in order to determine its inherent AC activity as well as further functionally characterising this enzymatic activity. The AtCAP-AC's ability to generate cAMP from ATP as

well as its different responses to the various molecules and chemical ions (ATP, GTP, Mg^{2+}, Mn^{2+}, Ca^{2+}, CO_3^{2-}, F^-) are presented below (Figure 5.3).

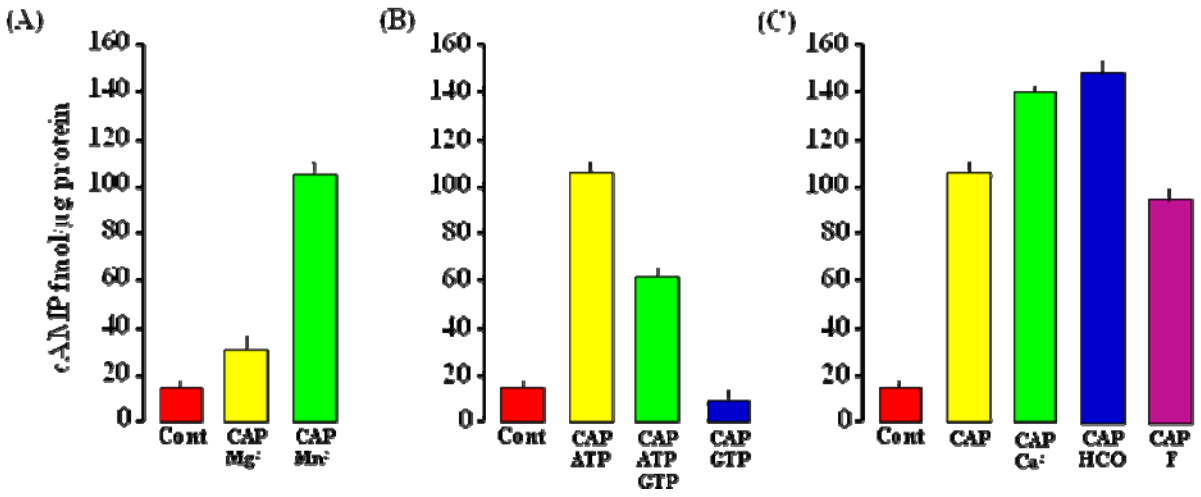

Figure 5.3: Molecular characterisation of the *in vitro* adenylate cyclase activity of the recombinant AtCAP-AC protein. A reaction mixture containing 25 µg of the purified recombinant AtCAP-AC protein, 50 mM Tris-HCl; pH: 8.0, 2 mM IBMX, 5 mM Mg^{2+}, 1 mM ATP and/or in the presence of other additives was incubated at room temperature for 20 minutes. The generated cAMP was then measured with a cAMP-specific enzyme immunoassaying kit (Catalogue # CA201; Sigma-Aldrich Inc., Missouri, USA) based on its acetylation protocol. (A) Cyclic AMP levels generated with the purified recombinant AtCAP-AC protein in the presence of magnesium or manganese ions, (B) cAMP levels generated with the purified recombinant AtCAP-AC in the presence of either ATP and/or GTP, and (C) cAMP levels generated with the purified recombinant AtCAP-AC protein in the presence of calcium, bicarbonate or fluoride ions. All error bars represent the standard errors of the means of triplicate assays.

5.4 DISCUSSION

The solubility/insolubility assessment undertaken in Section 5.2.1 demonstrated that the recombinant AtCAP-AC protein was wholly expressed in the form of an insoluble product and thus as inclusion bodies. This was probably a result of the host cells failing to tolerate the soluble nature of the foreign AtCAP-AC in their cytoplasms because of its excessive levels of expression (Rogl *et al*., 1998). Generally, the purification of recombinant proteins over-

expressed in *E. coli* systems is necessary so that all contaminating and toxic impurities are purged away (Tranken *et al*., 2000). In this study, an affinity purification process of the C-terminus 6x-His-tagged recombinant AtCAP-AC protein was successfully conducted using the Ni-NTA affinity system and under the non-native denaturing conditions (Figure 5.1). This is typically a fast and simple procedure which facilitates a tight association between the protein-fused tag and its corresponding affinity matrix, and thus allowing for the other protein contaminants to be easily washed off, even under the strongly and harsh denaturing conditions required to solubilise the inclusion bodies (Crowe *et al*., 1994). However and in order to carry out the desired subsequent *in vitro* enzyme immunoassaying system, the purified but denatured recombinant AtCAP-AC protein therefore, needed to be first reverted into its native, soluble and functional form, and through a well-designed refolding system (Singh and Panda, 2005) (Table 5.1), which in this study, was successfully achieved (Figure 5.2).

Generally, several studies have previously established magnesium as a necessary co-factor for most AC activities (Bradham, 1977). However, Steer and Levitzki (1975) also further found out that the Mn^{2+} ion could also successfully substitute Mg^{2+} ion as the co-factor activator of a turkey erythrocyte sAC. As a result of this, the recombinant AtCAP-AC activity was therefore hereby tested with regard to either or both of these two metal ions *in vitro*. Apparently, while the Mg^{2+} ion had no apparent effect onto the enzymatic activity of this recombinant protein, the Mn^{2+} ion significantly enhanced its enzymatic activity to a factor of approximately 7-fold (Figure 5.3A). Apparently, this finding on its own shows that even though both the Mg^{2+} and Mn^{2+} metal ions can be able to activate other known ACs in various organisms (Steer and Levitzki, 1975), the recombinant AtCAP-AC protein is strictly dependent onto the Mn^{2+} ion as its sole co-

factor for enzymatic activity. Nonetheless, this result is typically consistent with that of Braun and Dods, who in 1975, also showed that a sAC from the rat testis was strictly Mn^{2+}-dependent.

Subsequently, the recombinant AtCAP-AC protein was also further characterized with respect to its catalytic preference between the two structural analogs, ATP and GTP, as its sole substrates. The results showed that ATP alone could increase the levels of cAMP production with a factor of approximately 7-fold, while GTP alone drastically decreased the cAMP production by almost 33%, and on the other hand, the presence of GTP in an ATP reaction (ATP/GTP) had an apparent effect of decreasing the cAMP production with a factor of about 4-fold (Figure 5.3B). From this outcome, it is therefore very evident that even though the GTP could have had been structurally competing with ATP for the AtCAP-AC binding site, ATP is typically the preferred substrate for the AtCAP-AC protein.

The recombinant AtCAP-AC protein was once more, further characterized with respect to its catalytic responses in the presence of calcium. The results showed that the Ca^{2+} ion could significantly increase the catalytic activity of the AtCAP-AC protein by a factor of approximately 33% when ATP was the sole substrate (Figure 5.3C), and this outcome being very consistent with one of the previous findings which reported that sACs are typically stimulated by calcium (Jaiswal and Conti, 2003; Zippin *et al*., 2013). Concurringly, some previous studies on the effect of calcium onto the activity of sACs further suggested that the Ca^{2+} ion appears to affect these enzymes in a bi-phasic manner; i.e. stimulatory at low concentrations and inhibitory at high concentrations (Potter *et al*., 1980). Consistently, very relatively low concentrations of calcium were found to be necessary for the stimulation of the frog bladder epithelial sAC while all other concentrations higher than 0.5 mM were found to be rather inhibitory (Bockaert *et al*.,

1972). Moreover, the glioma sAC was found to be stimulated by less than or equal to 1 mM Ca^{2+} but inhibited as the concentration increased to approximately 100 mM (Brostrom *et al*., 1976). In our case and consistently, the recombinant AtCAP-AC protein was significantly stimulated by a 0.25 mM Ca^{2+} concentration, and therefore concurrently presenting it as a sAC. Interestingly, while all tmACs are mediated by the second messenger, cAMP via control mechanisms that are regulated by the GTP-binding proteins, all sACs are mediated by the second messenger, cAMP via control mechanisms that are regulated by the calcium-binding protein, calmodulin (Kamenetsky *et al*., 2006).

The activity of the recombinant AtCAP-AC protein was also further characterised with respect to the bicarbonate ion. The results showed that the bicarbonate ion could significantly increase the AtCAP-AC activity by a factor of about 40% when ATP was the sole substrate (Figure 5.3C). Previously, it has been shown that the membrane-bound (insoluble) AC of the fluid transporting tissue was stimulated by 10 mM bicarbonate and this activation was actually dose-dependent up to a concentration of around 100 mM (Mittag *et al*., 1993). Moreover, sACs can systematically synthesise cAMP in response to the bicarbonate ion (Steegborn *et al*., 2004), with the ion itself being able to activate the soluble mammalian AC up to 6-folds (Garty and Salomon, 1987), as well as activating the sperm sAC with a similar margin (Okamura *et al*., 1985). Furthermore, it has also been demonstrated that the bicarbonate ion could stimulate the mammalian sAC activities both *in vivo* and *in vitro* in a pH-independent manner (Chen *et al*., 2000); and a sAC from the cyanobacterium, *Anabaena* sp. PCC7120 was also stimulated by bicarbonate in this manner. Interestingly, it has further been reported that the sACs that contain the amino acid threonine in their active sites are bicarbonate-responsive, whereas those that contain a corresponding aspartate amino acid are bicarbonate-insensitive (Cann *et al*., 2003). Apparently,

the catalytic centre of our recombinant AtCAP-AC protein however, does not contain threonine, hence providing a potential evolutionary difference between plant sACs and the other general sACs in that they are stimulated by bicarbonate in the absence of an active site threonine. Cann *et al*. (2003) further asserted that the bicarbonate signalling through cAMP synthesis is a mechanism by which various organisms can now possibly respond to the ever-increasing environmental carbon.

Lastly, the AtCAP-AC protein fragment was characterised with respect to its modulation by the fluoride ion. The results showed that this ion has no significant effect onto the catalytic activity of the AtCAP-AC protein (Figure 5.3C). Previous studies by Braun *et al*. (1977) found a sAC from rat testis to be fluoride-insensitive and later on, Adamo *et al*. (1980) showed that a sAC from mouse testis was not stimulated with fluorine. Furthermore, a related study on a bull or boar sperm found out that sodium fluoride had no apparent effect onto the activity of a cytosolic sAC (Forte *et al*., 1983). In this study, the finding that the recombinant AtCAP-AC protein lacks sensitivity to the fluoride ion further bolsters our assertion that the AtCAP-AC is a sAC and it strongly concurs with our previous finding in Chapter 3, whereby the AtCAP-AC had also no catalytic responses to forskolin.

5.5 CONCLUSION

The findings of this chapter provide a solid and firm evidence that the recombinant AtCAP-AC protein is indeed a sAC whose activity is positively modulated by the manganese, calcium and bicarbonate ions while, on the other hand, it is also totally insensitive to forskolin, the magnesium, and fluoride ions. Therefore, we can now hereby firmly conclude that the AtCAP

protein from *Arabidopsis thaliana* is a *bona fide* sAC, whose physiological and biochemical activities may be mediated by cAMP and possibly, via a calmodulin-dependent signalling pathway.

5.6 RECOMMENDATIONS

There is still a great need to assess the functional activities of the recombinant AtCAP-AC protein in either detached living plant tissues or full plants so as to further augment our current wet bench experimental evidence with the exact *in vivo* and/or *in planta* activities of this promising protein. Other approaches such as mass spectrometry and/or the bioinformatic analysis of the AtCAP-AC protein across the whole Arabidopsis genome could be a better option to establish the typical signalling pathways in which this protein is involved and probably, leading to the possible complete elucidation of its exact physiological and/or functional roles in higher plants.

CHAPTER 6

BIOINFORMATIC EXPRESSIONAL ANALYSIS OF THE *AtCAP* GENE

ABSTRACT

The prototype second messenger molecule, adenosine 3′,5′-cyclic monophosphate (cAMP), generated by adenylate cyclases (ACs), is central to the numerous developmental and physiological processes of nearly all living organisms including plants. Of all the four ACs currently known in higher plants, the *Zea mays* pollen signalling protein is responsible for the polarized growth of pollen tubes, the *Arabidopsis thaliana* pentatricopeptide protein is involved in chloroplast biogenesis while the *Nicotiana benthamiana* and the *Hippeastrum hybridum* AC proteins have both been implicated in plant responses to biotic stress factors. Therefore, in an effort to attempt and elucidate the exact physiological roles of our cloned AtCAP-AC protein, we sequentially used a combination of various web-based computational tools to bioinformatically establish its functional roles and specifically, based on the patterns and trends of its expression profiles. Consequently, results from our findings indicated that the AtCAP-AC is typically involved in clathrin-coated vesicle-mediated biotic stress responses whose systems are principally mediated by cAMP and specifically via the SORLIP1AT regulatory core motif.

6.1 INTRODUCTION

Bioinformatics is increasingly becoming an indispensable tool for the applied research in biotechnology (Rehm, 2001). While in traditional biology, the typical research approach has been to study one gene at a time, in modern biotechnology, new technologies such as bioinformatics now allow researchers to study the genome-wide genes, and concurrently measure their changes and regulation under a variety of biological conditions (Huang *et al*., 2009).

Due to the increasing knowledge about plant metabolic pathways, it has become more apparent that the linear pathways previously studied are actually part of more complex networks with much overlap between their branches (Knight and Knight, 2001). Each gene is estimated to interact with about 4 to 8 other genes and to be involved in at least 10 different biological functions (D'haeseleer, 2000). Actually, the advances in genome sequencing have made it likely that there are numerous possibilities for intra-molecular and/or inter-molecular cross-talking among genes, a phenomenon which is increasingly becoming pertinent due to the bioinformatic identification and analysis of various plant genes and their protein products (Xiong *et al*., 2002). The reality that understanding a gene's network properties is nowadays, as important or more so, as understanding its function in isolation (Carlson *et al*., 2006).

Consequently, one of the most critical problems in making sense out of the sequenced genomes is the assignment of functional roles to the newly discovered proteins (Letovsky and Kasif, 2003). Gene sequencing has been able to identify the structure of thousands of genes, and the structural and regulatory features that provide functional clues (D'haeseleer, 2000), and if sequence similarity to one or more database sequences whose function is already known is obtained, then the unknown protein may possibly be inferred to have the same function,

biochemical activity or structure (Rehm, 2001). However, this notable approach of assigning protein function based on sequence similarity has since already failed for about 20-40% of proteins in the newly-sequenced genomes (Letovsky and Kasif, 2003). Even in the well-studied *Escherichia coli*, more than 30% of the open reading frames are currently hypothetical genes (Rehm, 2001). Moreover, in the model plant *Arabidopsis thaliana*, only about half of its genes have since been ascribed a function based on sequence similarity, and only about 11% have experimentally been tested for function (Usadel *et al*., 2009).

In this regard, it is essential to understand the whole network of the cellular context in order to comprehend gene function. It thus now appears that expression profiling and protein-protein interaction mapping provide additional opportunities for inferring protein function (Letovsky and Kasif, 2003). The correlation of gene expression patterns has since been used in plant science to make predictions about gene function and co-regulation, hence predicting their biology, exploring their system-level functionality and informing their experimental approaches (Zhang and Horvath, 2005; Manfield *et al*., 2006). Ideally, very large amounts of data can feasibly be generated through large-scale gene screening technologies such as mRNA hybridization micro-arrays. Microarray data provide information for the mRNA expression levels of thousands of genes in a tissue sample and have since been used, among other things, to identify genes associated with particular environmental stimuli and disease states (Jen *et al*., 2006). Microarray experiments are typically very powerful because they identify genes that are differentially regulated across various biological conditions (Choi and Kendziorski, 2009).

Knowledge of genome-wide gene expression patterns is normally required to understand the role of individual genes in complex biological processes that involve the interaction of many different

genes (Harr and Schlotterer, 2006). It appears that clusters of genes showing similar expression patterns (co-expression) across several experiments will tend to share biological functions and hence, functions of unknown genes may be hypothesized from genes with known function within the same cluster and their associated coherent patterns can then be accordingly modeled. In this respect, gene expression clustering is potentially useful in extracting regulatory motifs, in inferring functional annotations, and also as a molecular signature in distinguishing cell types (D'haeseleer *et al*., 2000; Jen *et al*., 2006; Meier and Donaldson, 2012). This co-expression data can then be complementary to information on possible gene function provided by sequence analysis (Jen *et al*., 2006) since measures of similarity in expression and function provide independent information about similarity in regulatory mechanisms (Allocco *et al*., 2004).

Generally, it is commonly accepted that in plants, genes which are functionally related may be transcriptionally coordinated. In the case of Arabidopsis, a significant amount of expression data from experiments on different tissues, developmental stages and stimuli are currently publicly available (Mutwil *et al*., 2008), meaning that the systems biological approach of depicting organization and functional relationships of component molecules can now be applied to this model plant (Aoki *et al*., 2007). For this reason, given that the genome of Arabidopsis has been fully sequenced, various web servers have been created based on diverse algorithms to measure the extents, similarities and patterns of gene expression under various conditions in the plant.

In this chapter, we therefore hereby detail the use of various web-based servers to bioinformatically elucidate the possible physiological roles of a clathrin assembly protein (AtCAP: At1g68110) from *Arabidopsis thaliana* based on the analysis of its expression patterns.

6.2 METHODOLOGY

In order to conduct the bioinformatic analysis of the *AtCAP* gene and in an effort to elucidate its possible and/or exact physiological roles in *Arabidopsis thaliana*, various bioinformatic web-based servers and programs were utilized.

6.2.1 THE EXPRESSION INTENSITY OF *AtCAP* IN ARABIDOPSIS TISSUES

The microarray database and expression-data analysis tool, GENEVESTIGATOR Version V3 (www.genevestigator.com/gv/) (Zimmermann *et al*., 2004; Grennan, 2006), was used to reveal the expression intensities of the *AtCAP* in 111 tissues of the *Arabidopsis thaliana*. This tool (GENEVESTIGATOR) was used because it provides the transcriptome information from the Affymetrix Arabidopsis ATH1 Genome Array platform using the 260011_At probe. Using At1g68110 as the query term, the arbitrary values of the expression intensity of the *AtCAP* in 111 Arabidopsis tissues were retrieved from the GENEVESTIGATOR.

6.2.2 CO-EXPRESSION ANALYSIS OF THE *AtCAP* IN *Arabidopsis thaliana*

The Arabidopsis co-expression tool, (ACT) (www.arabidopsis.leeds.ac.uk/ACT/), was used to establish the co-expression profile of the *AtCAP* and other related genes in the Arabidopsis. This web-based tool allows for the identification of genes whose expression patterns are correlated and based on an *Arabidopsis thaliana* microarray data set obtained from the Nottingham Arabidopsis Stock Centre (NASC). This ACT provides the correlation relationship results accompanied by estimates of their statistical significance, expressed as probability (P) and expectation (E) values (Jen *et al*., 2006). The co-expression analysis was performed across all available experiments using At1g68110 as the driver gene.

6.2.3 THE STIMULUS-SPECIFIC MICROARRAY EXPRESSION PROFILE OF THE *AtCAP-ECGG52*

Following the retrieval of the *AtCAP* co-expression group of 52 genes (*ECGG52*), the FatiGO+ tool in the Babelomics suite was then used to identify any significant enrichment in functional terms associated with the *AtCAP-ECGG52* (Al-Shahrour *et al*., 2007). Thereafter, using the stimulus tool, the expression profiles of the *AtCAP-ECGG52* were screened over the ATH1:22K array Affymetrix public microarray data in the GENEVESTIGATOR V3 version (www.genevestigator.com) (Zimmermann *et al*., 2004; Grennan, 2006). The normalized microarray data were downloaded from the GEO (NCBI) (www.ncbi.nlm.gov/geo/), the NASC Arrays (www.affymetrix.arabidopsis.info/narrays/experimentbrowse.pl) and the TAIR GenExpress (www.ebi.ac.uk/microarray-as/ac/) and subsequently analysed for experiments that were found to induce a differential expression of the genes. For each experiment found to induce differential expression, the fold-change (log2) values were then calculated. Subsequently, expression values were generated using the Multiple Array Viewer program from the Multi-Experiment Viewer (MeV) software package (Version 4.2.01) (The Institute for Genomic Research (TIGR)).

6.2.4 PROMOTER ANALYSIS OF THE *AtCAP-ECGG52*

The web-based visualisation tool in Athena (www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/visualize/pl) (O'Connor *et al*., 2005) analyses Arabidopsis promoter sequences and reports enrichments of known plant transcription factor binding sites (TFBSs). Enrichments in potential TFBSs were therefore analysed in the promoter regions of the *AtCAP* and the *AtCAP-ECGG52* using settings of 1 kb upstream of the predicted transcription start site (TSS)

and not cutting off at adjacent genes. The results of the Athena analysis were then confirmed in the POBO application (www.ekhidna.biocenter.jelsinki.ft/poxo/pobo/pobo) (Kankainen and Holm, 2004) by uploading promoter sequences 1 kb upstream of the coding regions of the *AtCAP-ECGG52*. The analysis was eventually run against a clean Arabidopsis background search for the SORLIP1AT core motif (GTGGC) using default settings. A 2-tailed p-value was then calculated in the linked online GraphPad website using the generated *t*-value and degrees of freedom to determine the statistical differences between the input sequences and its background.

6.3 RESULTS

6.3.1 The Expression Profile of *AtCAP* in Arabidopsis Tissues

The Genevestigator analysis of the *AtCAP* gene showed that the gene is significantly expressed in 111 plant tissues. However, there was a highly elevated expression level of this gene in the pollen and stamen of up to 5,700 arbitrary fluorescence units as compared to the rest of the other plant tissues whose expression levels were below 2,700 arbitrary fluorescence units (Figure 6.1).

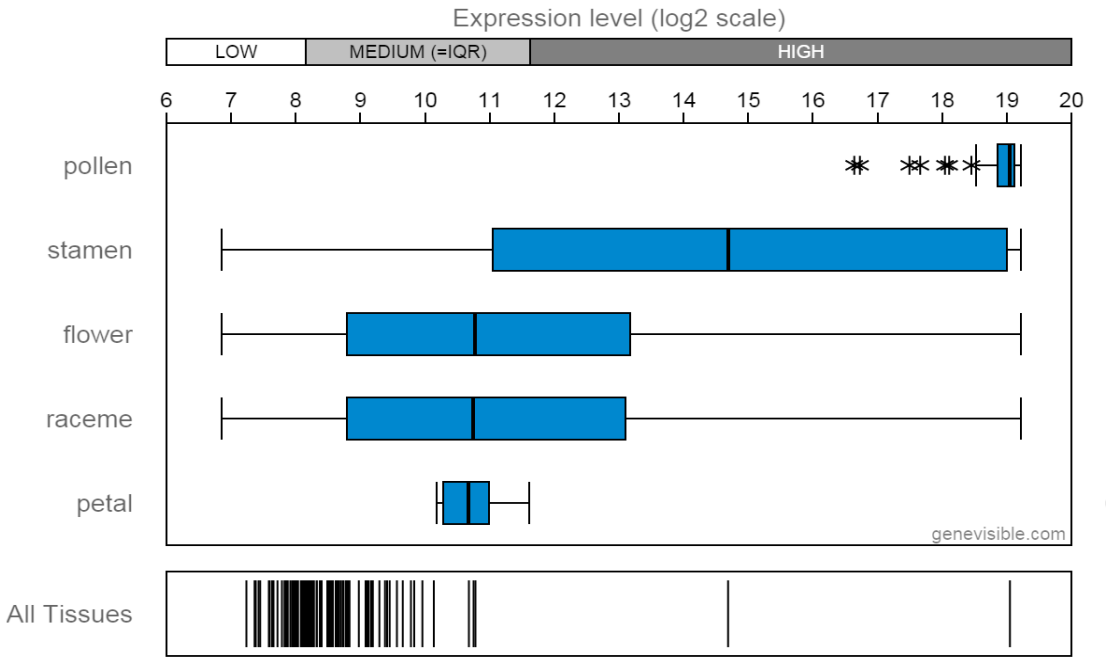

Figure 6.1: Expression intensity of the *AtCAP* in various tissues of the *Arabidopsis thaliana*. The *AtCAP* gene expression in various parts of the Arabidopsis plant showing considerably elevated expression levels in the pollen and stamen (data obtained from the Genevestigator anatomy tool (Zimmermann *et al*., 2004)).

6.3.2 Co-expression Analysis of the *AtCAP* in *Arabidopsis thaliana*

In order to understand the expression profile of the AtCAP gene in *Arabidopsis* thaliana, a desktop analysis of the publicly available microarray data was performed. Out of the 322 diverse transcriptome experiments, the AtCAP is highly co-expressed with numerous other genes in the Arabidopsis genome, with the top 52 genes having a Pearson correlation coefficient (r-value) of between 0.824 and 0.894 (Table 6.1). These 52 top most co-expressed genes (hereafter referred to as the AtCAP-Expression Correlated Gene Group (*AtCAP-ECGG52*)) were considered here because their correlation values were high and their number provided a moderately representative sample size for their subsequent stimulus-specific expression and promoter analysis evaluations.

Table 6.1: The top 52 genes whose expression profiles are directly correlated with that of *AtCAP* (At1g68110). All of these genes have Pearson correlation coefficients (r-value) between 0.824 and 0.894.

Rank	Locus and GO annotation	*R*	Description
1	AT1G68110[V-MT]	1	ENTH/ANTH/VHS superfamily protein
2	AT1G29970	0.894	60S ribosomal protein L18A-1
3	AT5G53360	0.89	TRAF-like superfamily protein
4	AT4G26400	0.886	RING/U-box superfamily protein
5	AT2G07180	0.884	Protein kinase superfamily protein
6	AT2G42780	0.871	INVOLVED IN: regulation of transcription.
7	AT1G15880[V-MT, PL, MF, PT]	0.866	Golgi snare 11
8	AT3G61160	0.865	Protein kinase superfamily protein
9	AT2G15240	0.860	UNC-50 family protein
10	AT5G51400	0.860	PLAC8 family protein
11	AT5G32440	0.857	Ubiquitin system component Cue protein
12	AT1G33250	0.856	Protein of unknown function (DUF604)
13	AT5G16110	0.855	Unknown protein; LOCATED IN: chloroplast.
14	AT2G46900	0.855	CONTAINS InterPro DOMAIN/s: Basic helix-loop-helix.
15	AT5G17290[PL, PT]	0.855	Autophagy protein Apg5 family
16	AT5G64920	0.852	COP1-interacting protein 8
17	AT3G07870	0.852	F-box and associated interaction domains-containing
18	AT1G06700	0.851	Protein kinase superfamily protein
19	AT2G02970	0.849	GDA1/CD39 nucleoside phosphatase family protein
20	AT1G15350	0.846	Unknown protein.
21	AT4G03030	0.842	Galactose oxidase/kelch repeat superfamily protein
22	AT1G33050	0.841	Unknown protein.
23	AT4G27880	0.840	Protein with RING/U-box and TRAF-like domains
24	AT2G36900[V-MT, PL, MF,PT]	0.840	Membrin 11
25	AT5G23670	0.839	Long chain base2
26	AT3G58670	0.839	Protein of unknown function (DUF1637)
27	AT3G10770	0.838	Single-stranded nucleic acid binding R3H protein
28	AT1G05840	0.838	Eukaryotic aspartyl protease family protein
29	AT4G08330	0.837	Unknown protein; LOCATED IN: plasma membrane.
30	AT4G22750	0.837	Sterol 4-alpha methyl oxidase 1-3
31	AT4G22753	0.837	DHHC-type zinc finger family protein
32	AT2G39100[PL, PT]	0.836	RING/U-box superfamily protein
33	AT2G26210	0.836	Ankyrin repeat family protein
34	AT3G02700	0.834	NC domain-containing protein-related
35	AT5G56750	0.834	N-MYC downregulated-like 1
36	AT2G29400	0.834	Type one protein phosphatase 1
37	AT1G26670[V-MT, PL, MF, PT]	0.833	Vesicle transport v-SNARE family protein
38	AT4G20380	0.833	LSD1 zinc finger family protein
39	AT1G15400	0.833	Unknown protein; LOCATED IN: plasma membrane.
40	AT4G17730[V-MT, PL, MF, PT]	0.832	Syntaxin of plants 23
41	AT5G16480	0.831	Phosphotyrosine protein phosphatases superfamily protein
42	AT3G12570	0.830	FYD
43	AT3G09320	0.830	DHHC-type zinc finger family protein
44	AT2G18280	0.829	Tubby like protein 2

45	AT3G01770	0.829	Bromodomain and extraterminal domain protein 10
46	AT5G20520	0.829	Alpha/beta-Hydrolases superfamily protein
47	AT3G11100	0.829	Sequence-specific DNA binding transcription factors
48	AT2G28370	0.828	Uncharacterised protein family (UPF0497)
49	AT5G46150	0.828	LEM3 (ligand-effect modulator 3) family protein / CDC50
50	AT4G32150[V-MT, PL, MF, PT]	0.827	Vesicle-associated membrane protein 711
51	AT2G20120	0.826	Protein of unknown function (DUF502)
52	AT2G20130	0.826	like COV 1]
53	AT3G12630	0.824	A20/AN1-like zinc finger family protein

Abbreviations for indicated GO terms:
V-MT= Vesicle-Mediated Transport (GO:0016192); PL = Protein Localization (GO:0008104); MF=Membrane Fusion (GO:0006944); PT= Protein Transport (GO: 0015031).

6.3.3 The Stimulus-specific Microarray Expression Profile of the *AtCAP-ECGG52*

In order to determine the biological processes in which the *AtCAP-ECGG52* is involved, a functional enrichment of this gene set was conducted using the FatiGO+ tool from the Babelomics suite (Al-Shahrour *et al*., 2007) and this analysis identified binding, transport and regulation of transcription as the possible molecular functions in which the *AtCAP-ECGG52* is primarily involved.

Furthermore and after establishing the co-expressional profile of the *AtCAP* gene and its associated *AtCAP-ECGG52* gene set, the *AtCAP* and its associated set (*AtCAP-ECGG52*) were also subjected to an *in silico* global expression analysis of the specific experimental conditions that were able to induce differential expression of both gene sets. The histogram generated from the microarray expression analysis revealed that the transcriptional processes of *AtCAP* and its associated *AtCAP-ECGG52* set are generally and collectively induced in response to a wide range of biotic stress factors (Figure 6.2).

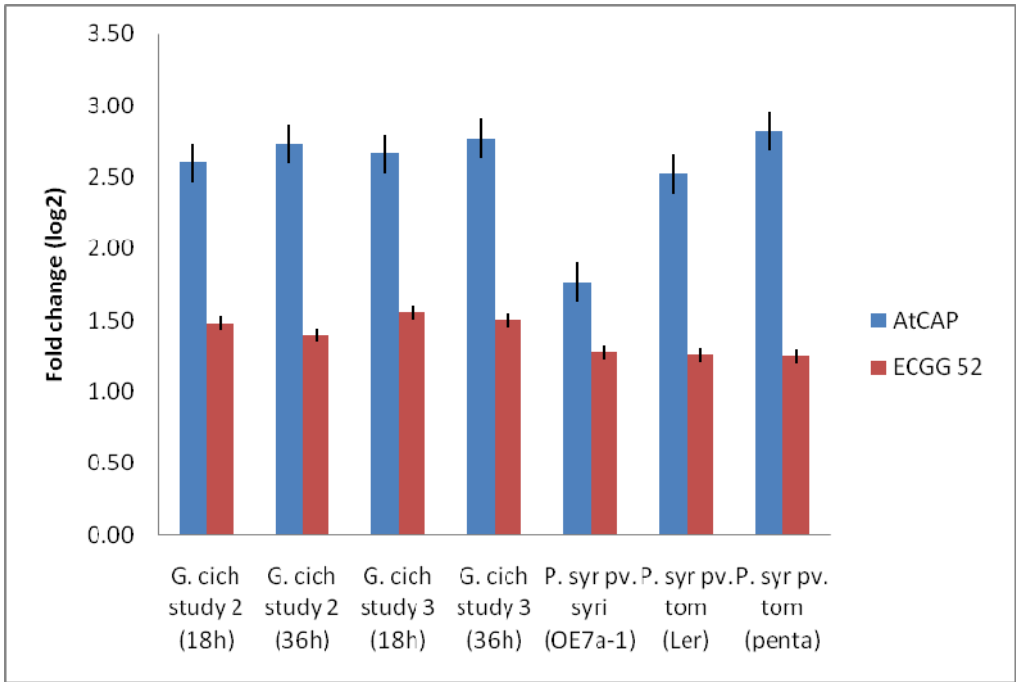

Figure 6.2: The expression profile of *AtCAP* and its *AtCAP-ECGG52* correlated genes in response to a selected array of biotic perturbations. The fold change in *AtCAP* expression of the *Arabidopsis thaliana* ecotype Columbia plants following (I) incubation with the pathogen elicitors *Golovinomyces cichoracearum* for eighteen hours and thirty six hours respectively (study 2); (II) expression of the edr (enhanced disease resistance) mutant plants following eighteen hours and thirty six hours respectively (study 3); (III) challenge with *Pseudomonas syringae* pv *syringae* after six hours treatment on OE7a-1 over-expression lines; (IV) infiltration with *Pseudomonas syringae* pv tomato virulent DC 3000 bacteria of 5-week old Ler plants collected twelve hours after treatment; and (V) infiltration of 5-week old penta mutant plants with *Pseudomonas syringae* pv tomato virulent DC 3000 bacteria collected after twelve hours infiltration.

6.3.4 Promoter Analysis of the AtCAP-ECGG52

The *AtCAP* and its associated *AtCAP-ECGG52* gene set display a significantly correlated expressional profile as well as a coordinated and differential expression response to specific biotic stress stimuli, which then suggests that these two gene sets may as well be under a common regulatory control and hence, may share common *cis*-regulatory elements in their promoter regions. Consequently, the promoter regions of the *AtCAP-ECGG5* were then analysed

for the presence of such suspected enriched *cis* elements using the Athena (O'Connor *et al*., 2005) and POBO (Kankainen and Holm, 2004) promoter analysis programs. The Athena analysis identified a significant enrichment in the SORLIP1AT core elements (GTGGC) and the subsequent POBO analysis (Kankainen and Holm, 2004) of the *AtCAP-ECGG52* then showed that the GTGGC motif was indeed significantly enriched (*t*-test; p-value >0.0001) (Figure 6.3).

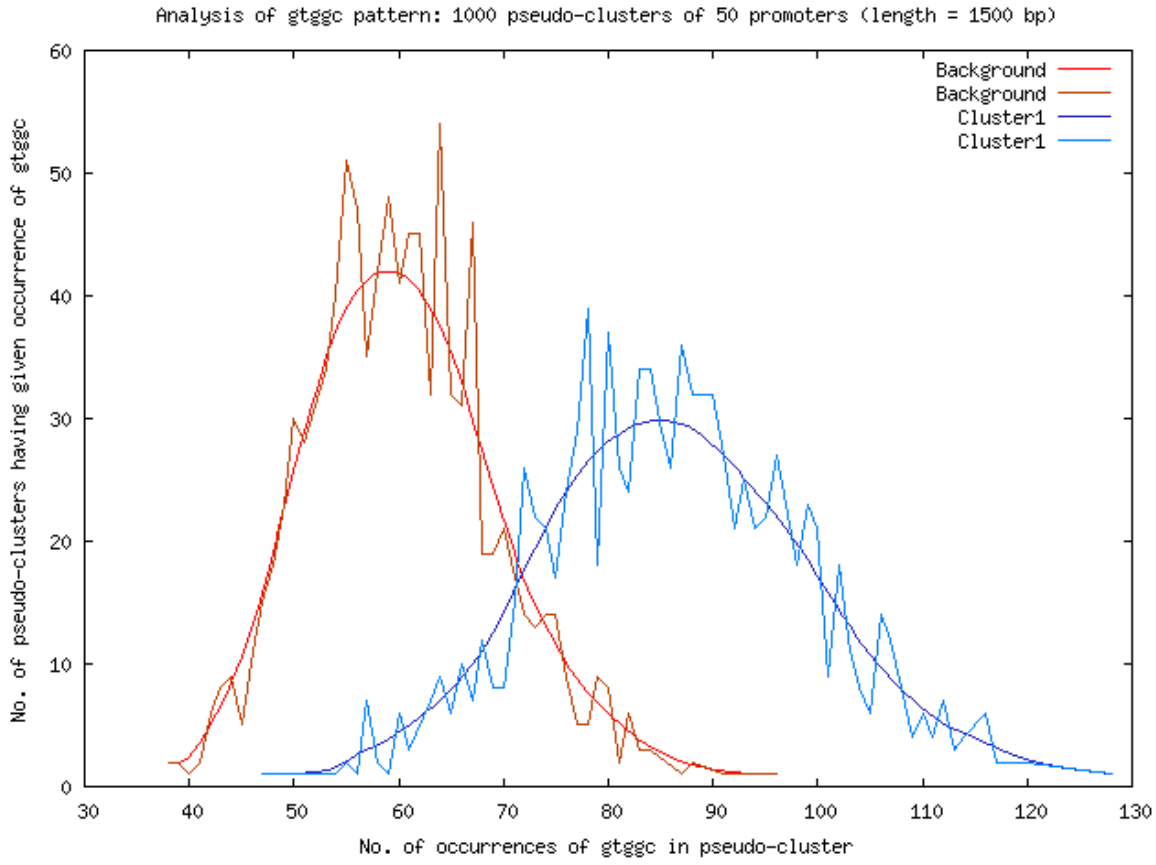

Figure 6.3: Frequency occurrence of the SORLIP1AT (GTGGC) motif in the promoters of the *AtCAP:ECGG52* and in the background *Arabidopsis thaliana* promoters. The 1-kb upstream promoter sequences of the *AtCAP-ECGG52* were analysed in the Athena (O'Connor *et al*., 2005) and POBO (Kankainen and Holm, 2004) programs to determine the frequency of occurrence of this SORLIP1AT GTGGC motif. The motif was then determined to be significantly enriched (*t*-test: p-value >0.001) in the *AtCAP-ECGG52*.

6.4 DISCUSSION

Results of the expressional analysis of the *AtCAP* gene in the Arabidopsis system indicate that this gene is principally expressed in the pollen grain and the stamen, and also significantly expressed in the flower, the raceme and the petal (Figure 6.1) alongside other genes that are specifically involved in transport and binding (Table 6.1). The *AtCAP* gene is also significantly up-regulated by various biotic stress perturbations, including some commonly known pathogenic elicitors like the *Golovinomyces cichoracearum* and the *Pseudomonas syringae* (Figure 6.2) and the mode of action of all these regulation appears to principally involve the SORLIP1AT *cis*-acting element with the GTGGC core motif (Figure 6.3).

These findings are significant because they correlate considerably with the other various experimental outcomes related to the cellular responses presented to both biotic and abiotic stress factors. Given that not all processes occurring in a plant are equally sensitive to stress, stress tolerance can therefore be developmentally and/or tissue-specifically regulated (Giorno *et al*., 2013). It has been reported that plant reproductive organs such as the pollen, the anthers and the flowers are highly sensitive to stress, as they strive to mitigate its negative impacts onto their sexual reproduction episodes (Giorno *et al*., 2013; Renak *et al*., 2014). For example, the sexual reproduction phase is considered as the most sensitive stage to environmental stress, and pollen specifically exhibiting the highest sensitivity during its development as stress at this stage can absolutely have negative and very deleterious developmental effects for meaningful and viable crop productive outcomes (Bokszczanin *et al*., 2013).

With regard to the biological processes in which the *AtCAP* is involved, binding and transport have both since been shown to have a central role in the mechanisms of stress response, and

particularly, the biotic stress responses. Binding has specifically been implicated in stress response (Parker *et al*., 2001; Shen *et al*., 2005; Shin *et al*., 2008), particularly metal binding, which has been linked to the removal of reactive oxygen species (ROS) and other toxic compounds that are normally generated during both the biotic and abiotic stress responses (Anantharaman *et al*., 2012; Zhang *et al*., 2014). Protein binding has also been linked to stress response whereby the proteins that mediate the arrest of growth and apoptosis of stressed cells are bound, thus supporting the survival of cells undergoing stress by desensitising them against such a stress (Hung *et al*., 2003; Flores-Kim and Darwin, 2015). In other cases, molecular chaperones, such as the heat-shock proteins involved in the environmental and disease-related stresses, assist newly-formed protein chains to fold correctly, thus preventing mis-foldings, which perhaps may lead to aggregations and eventual harmful inter-molecular interactions within the cell (Morimoto *et al*., 1997).

With regard to transport, it is essentially worthy to note that the *AtCAP* gene had previously been annotated as a clathrin assembly protein that binds to phosphatidylinositol (Berardini *et al*., 2004). Clathrin coat assembly drives the formation of clathrin-coated vesicles (Brodsky *et al*., 2001) which are responsible for a large portion of the vesicular trafficking that originates from the plasma membrane and the *trans*-Golgi apparatus network to endosomes (Schmid, 1997; Kirchhausen, 1999; Ehrlich *et al*., 2004). At cellular level, various molecules are internalized through clathrin-mediated endocytic pathways, hence clathrin-mediated endocytosis plays a crucial role in receptor activation, signal propagation and signal termination (Di Fiore and De Camilli, 2001). Endocytosis has thus been implicated in various stress response pathways which involve the phosphatidylinositol binding (Cavalli *et al*., 2001; Hayes *et al*., 2004; Mazel *et al*., 2004; Leshem *et al*., 2007). Hence, clathrin coats are involved in controlling the activity and

termination of signaling events, therefore apparently playing a key role into how the cell responds to its external environment (Le Roy and Wrana, 2005). Indeed, it has been asserted that, with regard to transport, it appears that several ubiquitous stress responses in *Arabidopsis thaliana*, commonly employ genes in pathways related to the vesicle transport (Ma and Bohnert, 2007), whereby various signal transduction pathways relay information via second messenger cascades through vesicle mediated trafficking systems (Seaman *et al*., 1996).

Naturally, plants usually respond to various stress factors through a direct and/or dramatic re-programming of their gene expressions by their respective transcription factors, leading to the relevant metabolic and phenotypic changes (Singh, 1998; Alves *et al*., 2014). As a result of this aspect, transcription factors are therefore recognized to be very important nuclear targets for cellular signal transduction systems. Notably, the SORLIP1AT (GTGGC) *cis*-element identified within 1 kb-upstream of the promoter region of the *AtCAP-ECGG52* is actually one of the Sequences Over-Represented in Light-Induced Promoters (SORLIPs) in Arabidopsis and is also one of the strand-independent computationally identified phytochrome A-induced motifs in the same plant (Hudson and Quail, 2005; Jiao *et al*., 2005; Corral *et al*., 2013). Typically, this motif is one of the *cis*-elements previously identified by Ma and Bohnert (2007) during the integration schemes of the Arabidopsis biotic and abiotic stress-related transcript profiles, promoter structures, and cell-specific expressions. Their finding has furthermore been firmly confirmed by several other subsequent studies which showed that the same SORLIP1AT motif is directly implicated in the regulation of gene expression, and mainly in response to the abiotic stress factors (Sharoni *et al*., 2011; Wang *et al*., 2011; Cakir and Giachino, 2012; Pegoraro *et al*., 2013; Chen *et al*., 2014; Yan *et al*., 2014; Zhao *et al*., 2014).

In this study therefore, we have shown that the *AtCAP* gene is generally up-regulated by the pathogenic elicitors and bacteria. This outcome is typically consistent with the results of some other previous studies where the Affymetrix whole-genome microarrays to profile the transcriptional regulation in Arabidopsis leaves after exposure to *Myzus persicae* aphid saliva as well as to aphid feeding showed that the *AtCAP* gene is one of the 78 genes whose transcriptional level was significantly induced (De Vos *et al*., 2005; De Vos and Jander, 2009). Thus and in light of all these findings, we therefore proposed that the AtCAP protein could principally be involved in a clathrin-coated vesicle-mediated stress response system that is mediated by the second messenger cAMP, and specifically regulated via the SORLIP1AT core motif.

6.5 CONCLUSION

Considering that the processes of transcription regulation and stress response in which the AtCAP protein is centrally involved require signal transduction mediated by second messengers, and that this study has experimentally demonstrated that the truncated adenylate cyclase domain of the AtCAP (AtCAP-AC) has the capacity to generate the second messenger molecule, cAMP, it is therefore apparent that the AtCAP protein is naturally, an important plant signaling molecule whose molecular actions are centrally mediated by cAMP and principally via the SORLIP1AT regulatory core motif.

6.6 RECOMMENDATION

In order to provide some *in planta* experimental evidence of the *AtCAP*'s exact AC biological activities and physiological roles in the plant's responses to the biotic stress factors, Arabidopsis

plants transformed with the *AtCAP-AC* gene can be biochemically assessed for changes in cAMP productions as well as for their genetic and phenotypic changes upon exposure to the different kinds of stress factors.

CHAPTER 7

General Discussion, Conclusion and Future Outlook

Since plants do play a very crucial and central role in the earth's ecosystem by providing food for consumers, their susceptibility both to biotic and abiotic stress factors poses a very huge challenge and threats to the sustainability of life on the planet. The major threats posed by climate change onto crop productivity have led scientists to develop innovative and creative interventions to try and improve the plants' abilities to withstand and/or adapt to these on-going stress factors. The most recent of these innovations, plant biotechnology, has proven to be very useful in the development of crop plants with desirable features and characteristics (Duvick, 2005; Yadav and Tyagi, 2006). Lately, the focus of plant biotechnology has specifically turned onto those plant molecules whose roles are centrally involved in homeostasis and signal transduction processes (Hussain *et al*., 2010). One such molecule is the cyclic adenosine 3′,5′-monophosphate (cAMP) generated by the enzyme, adenylate cyclase (AC) (Gasumov *et al*., 1999).

Apparently, information about higher plant ACs has, for a very long period of time, proved to be very elusive and a serious centre of controversy to the point that some scholars even went on to conclude that ACs do not at all exist in plants (Lomovatskaya *et al*., 2008). It was then suggested that until the enzymological characterization of a cloned or pure protein had been practically conducted, then the existence of higher plant ACs would always be a subject of very

critical and serious debates (Assmann, 1995). Consequently, the first ever higher plant AC to be experimentally and functionally confirmed was a *Zea mays* pollen signaling protein (PsiP) that was cloned and used to complement a *cya*A mutation in the *Escherichia coli* SP850 strain by Moutinho *et al*. in 2001. Thereafter, Gehring (2010) then bioinformatically identified 13 other putative AC candidates in the *Arabidopsis thaliana* using a search motif consisting of functionally assigned amino acids in the catalytic centers of annotated and/or experimentally tested nucleotide cyclases. Subsequently, Ruzvidzo *et al*. (2013) then experimentally proved that one of these putative Arabidopsis ACs, the pentatricopeptide protein (AtPPR), is actually a functional AC with a biological role in chloroplast biogenesis. Furthermore, two more ACs were eventually identified and experimentally confirmed in *Nicotiana benthamiana* (Ito *et al*., 2014) and *Hippeastrum hybridum* (Swiezawska *et al*., 2014) and were both shown be centrally involved in biotic stress responses. Consequently and under the premise that no single AC per plant could possibly be responsible for all the currently known cAMP-mediated processes in higher plants (Gehring, 2010), we then in this study, specifically focused onto attempting to experimentally prove the existence of yet another second AC candidate in the *A. thaliana*.

Firstly, we conducted a preliminary bioinformatic analysis, which then revealed that one of the twelve currently uncharacterized Arabidopsis putative ACs (Gehring, 2010), the clathrin assembly protein (AtCAP: At1g68110), is a multi-domain, multi-functional protein with a possible role in AC-dependent stress responses and adaptation mechanisms (Chapter 2). Subsequently, we then cloned the AC-containing fragment domain of this putative AtCAP protein (AtCAP-AC) and practically elucidated its exact biological mode of action. Upon a transient expression of this truncated AtCAP-AC recombinant in a chemically competent *Escherichia cloni* BL21 (DE3) pLysS DUOs prokaryotic system via a pTrcHis2-TOPO plasmid

expression vector, the protein induced a forskolin-insensitive generation of the endogenous cAMP molecule in this system, thereby exhibiting the property of being either a *bona fide* soluble AC (sAC) (Forte *et al*. 1983; Yan *et al*., 1998; Buck *et al*. 1999) capable of generating cAMP from ATP by itself or simply just yet another functional signaling plant molecule capable of stimulating other resident *E. cloni* ACs to produce their own cAMP (Chapter 3).

Thereafter, a molecular cloning and subsequent recombinant expression of the truncated AtCAP-AC cDNA into the mutant *Escherichia coli* strain (*cya*A SP850), which systematically lacks the AC activity, resulted in a functional and successful complementation of this mutant (Cotta *et al*., 1998; Charania *et al*. 2009) by the respectively expressed AtCAP-AC recombinant protein, which in turn succeeded to generate the then lacking but most required cAMP in this system. This finding on its own therefore, unequivocally confirmed the *in vivo* AC activity of the truncated AtCAP-AC protein, and thus firmly establishing it as a biologically functional AC - the fifth ever candidate to be confirmed in all higher plants and the second ever candidate to be confirmed in *A. thaliana* (Chapter 4).

Moreover, when the recombinant AtCAP-AC protein was eventually purified and functionally characterized *in vitro*, it demonstrated that its AC activity is strictly dependent on manganese and is positively modulated by both the calcium and carbonate ions, thus confirming itself as a *bona fide* soluble adenylate cyclase (sAC) (Potter *et al*., 1980; Okamura *et al*., 1985) whose physiological role/s may specifically be mediated by cAMP and possibly via a calmodulin-dependent signalling system (Chapter 5). Finally, a systematic bioinformatic assessment of the expressional profiles of this AtCAP-AC protein in the Arabidopsis plant (D'haeseleer *et al*., 2000; Jen *et al*., 2006; Meier and Donaldson, 2012) revealed that the protein is specifically

involved in biotic stress responses that are directly regulated by the SORLIP1AT expression core motif and via a cAMP-dependent signaling system (Chapter 6).

Hence by summing up all our findings in this study, we conclude that our work has firmly established the AtCAP as a functional sAC with specific roles in biotic stress responses and adaptation mechanisms in the *A. thaliana*. Additionally, we have also firmly provided evidence that the AtCAP protein is the fifth ever functional AC molecule to be experimentally identified in higher plants, after the PSiP from *Z. mays* (Moutinho *et al*., 2001), the AtPPR from *A. thaliana* (Ruzvidzo *et al*., 2013), the NbAC from *N. benthamiana* (Ito *et al*., 2014) and the HpAC1 from *H. hybridum* (Swiezawska *et al*., 2014). Furthermore, our findings also have positively indicated a likelihood that all the currently known and/or reported cAMP-dependent signalling processes in plants may possibly be mediated by more than one AC molecules since our studied AtCAP protein is the second ever such AC molecule to be experimentally identified and functionally confirmed in *A. thaliana*.

Finally, by considering the fact that the currently functionally characterized AtCAP protein had previously been bioinformatically identified together with the other twelve putative Arabidopsis AC candidates, it is recommended that the other outstanding eleven and functionally uncharacterized putative candidates be also functionally characterized, in order to gain a better understanding of their physiological and biological roles in plants, and particularly in critical cellular processes like growth, development and responses to stressful environmental factors. Furthermore, it is also highly recommended that more unknown AC candidates be enzymatically and functionally characterized in the agronomically important crop plants, with a view of elucidating their exact physiological and biological roles in stress response and adaptation

mechanisms, thus providing the basis for putting in place strategies and measures that enable plants to withstand the currently threatening environmental extremes brought about by climate change. Ideally, this could be an ultimate and valuable investment towards the possible eradication and/or sustainable alleviation of hunger and poverty on this troubled planet.

REFERENCES

Adamo S, Conti M, Geremia R and Monesi. **Particulate and soluble adenylate cyclase activities of mouse male germ cells.** *Biochemical and Biophysical Research Communications* 1980, **97(2):** 607-613.

AES Application Focus. **Gel electrophoresis of proteins.** AES Electrophoresis Society. Accessed at http://www.aesociety.org/areas/pdfs/Garfin_1DE_WebArticle9-07.pdf on 17/03/2015 at 19h20.

Ahle S and Ungewickell E. **Purification and properties of a new clathrin assembly protein.** *The European Molecular Biology Organization Journal* 1986, **5(2):** 3143-3149.

Ahle S, Mann A, Eichelsbacher U and Ungewickell E. **Structural relationships between clathrin assembly proteins from the Golgi and the plasma membrane.** *The European Molecular Biology Organization Journal Journal* 1988, **7(4):** 919-929.

Ahmad D and Newman EB. **A deficiency in cyclic AMP results in pH-sensitive growth of *Escherichia coli* K-12.** *Journal of Bacteriology* 1988, **170(8):** 3443-3447.

Allocco DJ, Kohane IS and Butte AJ. **Quantifying the relationship between co-expression, co-regulation and gene function.** *BioMedical Central Bioinformatics* 2004, **5:** 18.

Al-Shahrour F, Minguez P, Tarraga J, Medina I, Alloza E, Montaner D and Dopazo J. **FatiGO+: a functional profiling tool for genomic data. Integration of functional annotation, regulatory motifs and interaction data with microarray experimnents.** *Nucleic Acids Research* 2007, **35:** W91-W96.

Altschul SF, Gish W, Miller W, Myers EW and Lipman, DJ. **Basic local alignment search tool.** *Journal of Molecular Biology* 1990, **215:** 403-410.

Alves MS, Dadalto SP, Goncalves AB, de Souza GB, Barros VA and Fietto LG. **Transcription factor functional protein-protein interactions in plant defense responses.** *Proteomes* 2014, **2:** 85-106.

Anantharaman V, Iyer LM and Aravind L. **Ter-dependent stress response systems: novel pathways related to metal sensing, production of a nucleoside-like metabolite, and DNA-processing.** *Molecular Biosystems* 2012, **8(12):** 3142-3165.

Aoki K, Ogata Y and Shibata D. **Approaches for extracting practical information from gene co-expression networks in plant biology.** *Plant Cell Physiology* 2007, **48(3):** 381-390.

Assmann SM. **Cyclic AMP as a second messenger in higher plants: Status and future prospects.** *Plant Physiology* 1995, **108:** 885-889.

Bahzu O and Danchin A. **Adenylyl cyclases: A heterogeneous class of ATP-utilising enzymes.** *Progress in Nucleic Acid Research and Molecular Biology* 1994, **49:** 241-283.

Baker DA and Kelly JM. **Structure, function and evolution of microbial adenylyl and guanylyl cyclases.** *Molecular Microbiology* 2004, **52(5):** 1229-1242.

Baneyx F. **Recombinant protein expression in *Escherichia coli*.** *Current Opinion in Biotechnology* 1999, **10(5):** 411-421.

Beck KA and Keen JH. **Interaction of phosphoinositide cycle intermediates with the plasma membrane-associated clathrin assembly protein.** *The Journal of Biological Chemistry* 1991, **266:** 4442-4447.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J and Sayers EW. **GenBank.** *Nucleic Acids Research* 2013, **41(D1):** D36-D42.

Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G, Moseyko N, Yoo D, Xu I, Zoeckler B, Montoya M, Miller N, Weems D, Rhee SY. **Functional annotation of the Arabidopsis genome using controlled vocabularies.** *Plant Physiology* 2004, **135(2):** 745-755.

Blanco E, Parra G and Guigo R. **Using geneID to identify genes.** *Current Protocols in Bioinformatics* 2007, **18(4.3):** 4.3.1-4.3.2.

Bockaert J, Roy C and Jard S. **Oxytocin-sensitive adenylate cyclase in frog bladder epithelial cells.** *The Journal of Biological Chemistry* 1972, **247:** 7073-7081.

Bokszczanin KL, Solanaceae Pollen Thermotolerance Initial Training Network (SPOT-ITN) Consortium and Fragkostefanakis S. **Perspectives on deciphering mechanisms underlying plant heat stress response and thermo-tolerance.** *Frontiers in Plant Science* 2013, **4(315):** 1-20.

Bradham LS. **Fluoride activation of rat brain adenylate cyclase: the requirement for a protein co-factor.** *Journal of Cyclic Nucleotide Research* 1977, **3(2):** 119-128.

Braun T and Dods RF. **Development of a Mn^{2+}-sensitive, "soluble" adenylate cyclase in rat testis.** *Proceedings of National Academy of Sciences* 1975, **72(3):** 1097-1101.

Braun T, Frank H, Dods R and Sepsenwol S. **Mn^{2+} sensitive, soluble adenylate cyclise in rat testis: Differentiation from other testicular nucleotide cyclases.** *Biochemical and Biophysical Research Communication* 1977, **481(1):** 227-235.

Brodsky FM, Chen CY, Knuehl C, Towler MC and Wakeham DE. **Biological basket weaving: Formation and function of clathrin-coated vesicles.** *Annual Review of Cell and Developmental Biology* 2001, **17:** 517-568.

Brostrom MA, Brostrom CO, Breckenridge BM and Wolff DJ. **Regulation of adenylate cyclase from glial tumor cells by calcium and a calcium-binding protein.** *The Journal of Biological Chemistry* 1976, **251:** 4744-4750.

Buchan DWA, Minneci F, Nugent TCO, Bryson K and Jones DT. **Scalable web services for the PSIPRED Protein Analysis Workbench.** *Nucleic Acids Research* 2013, **41(W1):** W340-W348.

Buck J, Meeghan SL, Schapal L, Cann MJ and Levin LR. **Cytosolic adenylyl cyclase defines a unique signaling molecule in mammals.** *Proceedings of the National Academy of Sciences USA* 1999, **96(1):** 79-84.

Cakir B and Giachino RRA. ***VvTMT2* encodes a putative tonoplast monosaccharide transporter expressed during grape (*Vitis vinifera* cv. Sultanine) ripening.** *Plant Omics Journal* 2012, **5(6):** 576-583.

Cann MJ, Hammer A, Zhou J and Kanacher T. **A defined subset of adenylyl cyclases is regulated by bicarbonate ion.** *The Journal of Biological Chemistry* 2003, **278:** 35033-35038.

Carlson MRJ, Zhang B, Fang Z, Mischel PS, Horvath S and Nelson SF. **Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks.** *BioMedical Central Genomics* 2006, **7:** 40.

Carolan M. ***The sociology of food and agriculture.*** 2012. Oxford, UK: Routledge.

Cavalli V, Vilbois F, Corti M, Marcot MJ, Tamura K, Karin M, Arkinstall S and Gruenberg J. **The stress-induced MAP kinase p38 regulates endocytic trafficking via the GDI:Rab5 complex.** *Molecular Cell* 2001, **7(2):** 421-432.

Cech SY, Broaddus WC and Maguire ME. **Adenylate cyclase: The role of magnesium and other divalent cations.** *Molecular and Cellular Biochemistry* 1980, **33(1-2):** 67-92.

Chang YC, Hsiao YM, Wu MF, Ou CC, Lin YW, Lue KH and Ko JL. **Interruption of lung cancer cell migration and proliferation by fungal immunomodulatory protein FIP-fve from *Flammulina velutipes*.** *Journal of Agricultural and Food Chemistry* 2013, **61(49):** 12044-12052.

Charania MA, Brockman KL, Zhang Y, Banerjee A, Pinchuk GE, Fredrickson JK, Beliaev AS and Saffarini DA. **Involvement of a membrane-bound class III adenylate cyclase in regulation of anaerobic respiration in *Shewanella oneidensis* MR-1.** *Journal of Bacteriology* 2009, **191(13):** 4298-4306.

Chawla HS. ***Introduction to plant biotechnology. 2nd Edition.*** 2002. New Hampshire, USA: Science Publishers Inc.

Chen BY and Janes HW. ***PCR Cloning Protocols.*** 2002. New Jersey, USA: Humana Press.

Chen F, Li B, Demone J. Charron JB, Shi X and Deng XW. **Photoreceptor partner FHY1 has an independent role in gene modulation and plant development under far-red light.** *Proceedings of the National Academy of Sciences USA* 2014, **111(32):** 11888-11893.

Chen Y, Cann MJ, Litvin TN, Iourgenko V, Sinclair ML, Levin LR and Buck J. **Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor.** *Science* 2000, **289(5479):** 625-628.

Choi Y and Kendziorski C. **Statistical methods for gene set co-expression analysis.** *Bioinformatics* 2009, **25(21):** 2780-2786.

Clark DP and Pazdernik NJ. ***Biotechnology: Academic Cell Update Edition.*** 2011. London, UK Academic Press.

Cooper DMF. **Molecular and cellular requirements for the regulation of adenylate cyclases by calcium.** *Biochemical Society Transactions* 2003, **31(5):** 912.

Corral JM, Vogel H, Aliyu M, Hensel G, Thiel T, Kumlehn J and Sharbel TF. **A conserved apomixes-specific polymorphism is correlated with exclusive exonuclease expression in premeiotic ovules of apomictic *Boechera* species.** *Plant Physiology* 2013, **163(4):** 1660-1672.

Côté R, Reisinger F, Martens L, Barsnes H, Vizcaino JA and Hermjakob H. **The Ontology Lookup Service: bigger and better.** *Nucleic Acids Research* 2010, **38(Suppl 2):** W155-160.

Cotta MA, Whitehead TR and Wheeler MB. **Identification of a novel adenylate cyclase in the ruminal anaerobe, *Prevotella ruminicola* D31d.** *Federation of European Microbiological Societies - Microbiology Letters* 1998, **164(2):** 257-260.

Crowe J, Dobeli H, Gentz R, Hochuli E, Stuber D and Henco K. **6xHis-Ni-NTA chromatography as a superior technique in recombinant protein expression/purification.** *Protocols for Gene Analysis* 1994, **31:** 371.

Crowe J, Masone BS and Ribbe J. **One-step purification of recombinant proteins with the 6xHis tag and Ni-NTA resin.** *Basic DNA and RNA Protocols* 1996, **58:** 491-510.

D'haeseleer P, Liang S and Somogyi R. **Genetic network inference: from co-expression clustering to reverse engineering.** *Bioinformatics* 2000, **16(8):** 707-726.

Davies PJ. ***Plant hormones biosynthesis, signal transduction, action!*** 2004. Dordrecht, Netherlands: Springer Science and Business Media.

De Craene JO, Ripp R, Lecompte O, Thompson J, Poch O and Friant S. **Evolutionary analysis of the ENTH/ANTH/VHS protein superfamily reveals a coevolution between membrane trafficking and metabolism.** *BioMedical Central Genomics* 2012, **13(1):** 297.

De Gunzburg J. **Mode of action of cyclic AMP in prokaryotes and eukaryotes, CAP and cAMP-dependent protein kinases.** *Biochemistry* 1985, **67(6):** 563-582.

De Vos M and Jander G. ***Myzus persicae* (green peach aphid) salivary components induce defence responses in *Arabidopsis thaliana*.** *Plant, Cell and Environment* 2009, **32(11):** 1548-1560.

De Vos M, Van Oosten RM, Van Poecke RMP, Van Pelt JA, Pozo MJ, Mueller MJ, Buchala AJ, Metraux JP, Van Loon LC, Dicke M and Pieterse CMJ. **Signal signature and transcriptome changes of Arabidopsis during pathogen and insect attack.** *Molecular Plant Microbe Interactions* 2005, **18:** 923-937.

Dessler AF. ***Introduction to modern climate change.*** 2012. New York, USA: Cambridge University Press.

Di Fiore PP and De Camilli P. **Endocytosis and signaling: An inseparable partnership.** *Cell* 2001, **106(1):** 1-4.

Diaz ET, Lennarson R, Richard R, Bagajewicz M, and Harrison RG. **Prediction of Protein Solubility in *Escherichia coli* Using Logistic Regression.** *Biotechnology and Bioengineering* 2009, **105(2):** 374-383.

Dole S, Klingen Y, Nagarajavel V and Schnetz K. **The protease Lon and the RNA-binding protein Hfq reduce silencing of the *Escherichia coli* bgl operon by H-NS.** *Journal of Bacteriology* 2004, **186(9):** 2708-2716.

Duvick DN. ***Genetic and production innovations in field crop technology: New developments in theory and practice.*** 2005. New York, USA: Food Products Press.

Ehrlich M, Boll W, van Oijen A, Hariharan R, Chandran K, Nibert ML and Kirchhausen T. **Endocytosis by random initiation and stabilization of clathrin-coated pits.** *Cell* 2004, **118(5):** 591-605.

Elcock AH. **Prediction of functionally important residues based solely on the computed energetic of protein structure.** *Journal of Molecular Biology* 2001, **312(4):** 885-896.

Engqvist-Goldstein AEY, Warren RA, Kessels MM, Keen JH, Heuser J and Drubin DG. **The actin-binding protein Hi1R associates with clathrin during early stages of endocytosis and promotes clathrin assembly *in vitro*.** *Journal of Cell Biology* 2001, **154(6):** 1209-1224.

Fain JN, Pointer RH and Ward WF. **Effects of adenosine nucleosides on adenylate cyclase, phosphodiesterase, cyclic adenosine monophosphate accumulation, and lipolysis in fat cells.** *The Journal of Biological Chemistry* 1972, **247:** 6866-6872.

FAO. **How to feed the world in 2050.** A Food and Agriculture Organization Expert Paper. Accessed at http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf on 24/04/2015 at 20h54.

Farrokhi N, Hrmova M, Burton RA and Fincher GB. **Heterologous and cell-free protein expression systems.** *Plant Genomics in Methods in Molecular Biology* 2009, **513:** 175-198.

Fido RJ, Mills ENC, Rigby NM and Shewry PR. ***Protein Extraction from plant tissues in Protein Purification Protocols.*** 2004. New Jersey, USA: Humana Press Inc.

Flores-Kim J and Darwin AJ. **Activity of a bacterial cell envelope stress response is controlled by the interaction of a protein-binding domain with different partners.** *Journal of Biological Chemistry* 2015, **13:** 75-98.

Florio VA and Ross EM. **Regulation of the catalytic component of adenylate cyclase. Potentiative interaction of stimulatory ligands and 2′,5′-dideoxyadenosine.** *Molecular Pharmacology* 1983, **24(2):** 195-202.

Forte LR, Bylund DB and Zahler WL. **Forskolin does not activate sperm adenylate cyclase.** *Molecular Pharmacology* 1983, **24(1):** 42-47.

Francis DM and Page R. ***Strategies to optimize protein expression in E. coli.*** Wiley Online Library. Accessed at http://www.currentprotocols.com/WileyCDA/CPUnit/refId-ps0524.html on 22/01/2015 at 15h37.

Frommer WB and Ninnemann O. **Heterologous expression of genes in bacterial, fungal, animal, and plant cells.** *Annual Review of Plant Physiology and Plant Molecular Biology* 1995, **46:** 419-444.

Garty NB and Salomon Y. **Stimulation of partially purified adenylate cyclase from bull sperm by bicarbonate.** *Federation of the European Biochemical Societies Letters* 1987, **218(1):** 148-152.

Gasumov KG, Shichijo C, Bayramov SM and Hashimoto T. **Membrane and soluble fractions of adenylyl cyclase from *Sorghum bicolor* seedlings positively react to the action of red and far red lights.** *Internet Photochemistry and Photobiology.* Accessed at http://www.photobiology.com/index.htm on 14/03/2013 at 08h52.

Gavrilescu M and Christi Y. **Biotechnology – a sustainable alternative for chemical industry.** *Biotechnology Advances* 2005, **23(7):** 471-499.

Gehring C. **Adenylyl cyclases and cAMP in plant signaling – past and present.** *Cell Communication and Signalling* 2010, **8:** 15.

Giorno F, Wolters-Arts M, Mariani C and Rieu I. **Ensuring reproduction at high temperatures: The heat stress response during anther and pollen development.** *Plants* 2013, **2(3):** 489-506.

Gomperts BD, Kramer IM and Tatham PER. ***Signal Transduction (Second Edition).*** 2009. Massachusetts, USA: Academic Press.

Grennan AK. **Genevestigator: Facilitating web-based gene-expression analysis.** *Plant Physiology* 2006, **141(4):** 1164-1166.

Grodberg J and Dunn JJ. ***ompT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification.** *Journal of Bacteriology* 1988, **170:** 1245-1253.

Gross MK, Douglas CA, Smith AL and Storm DR. **Targeted mutations that ablate either the adenylate cyclase or haemolysin function of the bifunctionalcyaA toxin of *Bordetella***

***pertussis* abolish virulence.** *Proceedings of the National Academy of Science USA* 1992, **89:** 4898-4902.

Grossman TH, Kawasaki ES, Punreddy SR and Osburne MS. **Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability.** *Gene* 1998, **209(1-2):** 95-103.

Gurder SK. ***Scientific breakthroughs and plant breeding.*** 2005. New York, USA: Food Products Press.

Gustafsson C, Govindarajan S and Minshull J. **Codon bias and heterologous protein expression.** *Trends in Biotechnology* 2004, **22(7):** 346-353.

Hames D and Hooper N. ***Biochemistry: Instant Notes.*** 2000. Oxford, UK: Taylor and Francis.

Hanoune J and Defer N. **Regulation and role of adenylyl cyclase isoforms.** *Annual Review of Pharmacology and Toxicology* 2001, **41:** 145-174.

Harr B and Schlotterer C. **Comparison of algorithms for the analysis of Affymetrix microarray data as evaluated by co-expression of genes in known operons.** *Nucleic Acids* Research 2006, **34(2):** e8.

Haslam RJ, Davidson MM and Desjardins JV. **Inhibition of adenylate cyclase by adenosine analogues in preparations of broken and intact human platelets: Evidence for the unidirectional control of platelet function of cyclic AMP.** *Biochemical Journal* 1978, **176:** 83-95.

Hayes MJ, Merrifiels CJ, Shao D, Ayala-Sanmartin J, d'SouzaSchorey C, Levine TP, Proust J, Curran J, Bailly M and Moss SE. **Annexin 2 binding to phosphatidylinositol 4,5-biphosphate on endocytic vesicles is regulated by the stress response pathway.** *The Journal of Biological Chemistry* 2004, **279:** 14157-14164.

Heazlewood JL, Verboom RE, Tonti-Filippini J, Small I and Millar AH. **SUBA3: The Arabidopsis subcellular database.** *Nucleic Acids Research* 2013, **35(D):** 213-218.

Hess KC, Jones BH, Marquez B, Chen Y, Ord TS, Kamenetsky M, Miyamoto C, Zippin JH, Kopf GS, Suarez SS, Levin LR, Williams CJ, Buck J and Moss SB. **The "soluble" adenylyl cyclase in sperm mediates multiple signaling events required for fertilization.** *Developmental Cell* 2005, **9(2):** 249-259.

Hintermann R and Parish RW. **Determination of adenylate cyclase activity in a variety of organisms: Evidence against the occurrence of the enzyme in higher plants.** *Plant* 1999, **146:** 459-461.

Hochuli E, Dobeli H and Schacher A. **New chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues.** *Journal of Chromatography* 1987, **411:** 177-184.

Hogeweg P. **The roots of bioinformatics in theoretical biology.** *PLoS Computational Biology* 2011, **7(3):** e1002021.

Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem,W and Zimmermann P. **Genevestigator V3: A reference expression database for the meta-analysis of transcriptomes.** *Advances in Bioinformatics* 2008, **2008(420747):** 1-5.

Huang DW, Sherman BT and Lempicki RA. **Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.** *Nucleic Acids Research* 2009, **37(1):** 1-13.

Hudson ME and Quail PH. **Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data.** *Plant Physiology* 2003, **133:** 1606-1616.

Hung WJ, Roberson RS, Taft J and Wu DY. **Human BAG-1 proteins bind to the cellular stress response protein GADD34 and interfere with GADD34 functions.** *Molecular Cell Biology* 2003, **23(10):** 3477-3486.

Hussain K, Nisar MF, Majeed A, Nawaz K, Bhatti KH, Afghan S, Shahazad A and Zia-ul-Hussnian S. **What molecular mechanism is adapted by plants during salt stress tolerance?** *African Journal of Biotechnology* 2010, **9(4):** 416-422.

Hyne RV and Garbers DL. **Regulation of guinea pig sperm adenylate cyclase by calcium.** *Biology of Reproduction* 1979, **21(5):** 1135-1142.

Ichikawa T, Suzuki Y, Czaja I, Schommer C, LeBnick A, Schell J and Walden R. **Identification and role of adenylyl cyclase in auxin signaling in higher plants.** *Nature* 1997, **390:** 698-701.

Ito M, Takahashi H, Sawasaki T, Kouhei O, Hikichi Y and Kiba A. **Novel type of adenylyl cyclase participates in tabtoxinine-β-lactam-induced cell death and occurrence of wildfire disease in *Nicotiana benthamiana*.** *Plant Signaling and Behavior* 2014, **9(1):** e27420.

Jaiswal BS and Conti M. **Calcium regulation of the soluble adenylyl cyclase expressed in mammalian spermatozoa.** *Proceedings of the National Academy of Sciences USA* 2003, **100(19):** 10676-10681.

Jen CH, Manfield IW, Michalopoulos I, Pinney JW, Willats WGT, Gilmartin PM and Westhead DR. **The Arabidopsis co-expression tool (ACT): a WWW-based tool and database for microarray-based gene expression analysis.** *The Plant Journal* 2006, **46:** 336-348.

Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P and von Mering C. **STRING8 – a global view on proteins and their functional interactions in 630 organisms.** *Nucleic Acids Research* 2009, **37:** D4120416.

Jiao Y, Ma L, Strickland E and Deng XW. **Conservation and divergence of light-regulated genome expression patterns during seedling development in rice and Arabidopsis.** *Plant Cell* 2005, **17:** 3239-3256.

Kamenetsky M, Middelhaufe S, Bank EM, Levin LR, Buck J and Steegborn C. **Molecular details of cAMP generation in mammalian cells: A tale of two systems.** *Journal of Molecular Biology* 2006, **362(4):** 623-639.

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M and Tanabe M. **Data, information, knowledge and principle: back to metabolism in KEGG.** *Nucleic Acids Research* 2014, **42:** D199-D205.

Kanehisa M. **Linking databases and organisms: GenomeNet resources in Japan.** *Trends in Biochemical Sciences* 1997, **22(11):** 442-444.

Kankainen M and Holm L. **POBO, transcription factor binding site verification with bootstrapping.** *Nucleic Acids Research* 2004, **32:** W222-W229.

Kast EJ, Nguyen MDT, Lawrence RE, Rabeler C and Kaplinsky NJ. **The RootScope: a simple high-throughput screening system for quantitating gene expression dynamics in plant roots.** *BioMedical Central Plant Biology* 2013, **13:** 158.

Katsumata T, Takahasi N and Ejiri S. **Changes of cyclic AMP level and adenylate cyclase activity during germination of pine polled.** *Agricultural and Biological Chemistry* 1978, **42(11):** 2161-2162.

Keen JH. **Clathrin assembly proteins: Affinity purification and a model for coat assembly.** *The Journal of Cell Biology* 1987, **105:** 1989-1998.

Kelley LA and Sternberg MJE. **Protein structure prediction on the web: a case study using the Phyre Server.** *Nature Protocols* 2009, **4:** 363-371.

Kersey PJ, Allen JE, Christensen M, Davis P, Falin LJ, Grabmueller C, Hughes DST, Humphrey J, Kerhornou A, Khobova J, Langridge N, McDowall MD, Maheswari U, Maslen G, Nuhn M, Ong CK, Paulini M, Pedro H, Toneva I, Tuli MA, Walts B, Williams G, Wilson D, Youens-Clark K, Monaco MK, Stein J, Wei X, Ware D, Bolser DM, Howe KL, Kulesha E, Lawson D and Staines DM. **Ensembl Genomes 2013: scaling up access to genome-wide data.** *Nucleic Acids Research* 2004, **42(D1):** D546-D552.

Kilikian BV, Suarez ID, Liria CW and Gombert AK. **Process strategies to improve heterologous protein production in *Escherichia coli* under lactose or IPTG induction.** *Process Biochemistry* 2000, **35(9):** 1019-1025.

Kim DY, Hong MJ, Lee YJ lee MB and Seo YW. **Wheat truncated hemoglobin interacts with photosystem I PSK-I subunit and photosystem II subunit PsbS1.** *Plant Biology* 2013, **57(2):** 281-290.

Kirchhausen T. **Adaptors for clathrin-mediated traffic.** *Annual Review of Cell and Developmental Biology* 1999, **15:** 705-732.

Knight H and Knight MR. **Abiotic stress signaling pathways: specificity and cross-talk.** *Trends in Plant Science* 2001, **6(6):** 262-267.

Kok JN. ***Knowledge Discovery in Databases: PKDD 2007.*** 2007. New York, USA: Springer Science and Business Media.

Koshiba S, Kigawa T, Kikuchi A and Yokoyama S. **Solution structure of the epsin N-terminal homology (ENTH) domain of human epsin.** *Journal of Structural and Functional Genomics* 2001, **2:** 1-8.

Krauss G. ***Biochemistry of Signal Transduction and Regulation.*** 2008. Weinheim, Germany: John Wiley and Sons.

Krishnakumar V, Hanlon MR, Contrino S, Ferlanti ES, Karamycheva S, Kim M, Rosen BD, Cheng CY, Moreira W, Mock SA, Stubbs J, Sullivan JM, Krampis K, Miller JR, Micklem G, Vaughn M and Town CD. **Araport: the Arabidopsis information portal.** *Nucleic Acids Research* 2015, **43(D1):** D1003-D1009.

Kuo JT, Chang YJ and Tseng CP. **Growth rate regulation of lac operon expression in *Escherichia coli* is cyclic AMP dependent.** *Federation of the European Biochemical Societies Letters* 2003, **553(3):** 397-402.

Kwezi L, Meier S, Mungur L, Ruzvidzo O and Irving H. **The *Arabidopsis thaliana* Brassinosteroid receptor (AtBRI1) contains a domain that functions as a guanylyl cyclase *in vitro*.** *PLoS ONE* 2007, **2(5):** e449.

Kwezi L, Ruzvidzo O, Wheeler JI, Govender K, Iacuone S, Thompson PE, Gehring C and Irving HR. **The Phytosulfokine (PSK) receptor is capable of guanylate cyclase activity and enabling cyclic GMP-dependent signalling in plants.** *The Journal of Biological Chemistry* 2011, **286:** 22580-22588.

Ladilov Y and Appukuttan A. **Role of soluble adenylyl cyclase in cell death and growth.** *Biochemical and Biophysical Research Communication – Molecular Basis of Disease* 2014, **1842(12):** 2646-2655.

Lalli E and Sassone-Corsi P. **Signal transduction and gene regulation: the nuclear response to cAMP.** *The Journal of Biological Chemistry* 1994, **269(26):** 17359-17362.

Laurenza A, Sutkowski EH and Seamon KB. **Forskolin: a specific stimulator of adenylyl cyclase or a diterpene with multiple sites of action?** *Trends in Pharmacological Sciences* 1989, **10(11):** 442-447.

Le Roy C and Wrana JL. **Clathrin- and non-clathrin-mediated endocytic regulation of cell signaling.** *Nature Reviews Molecular Cell Biology* 2005, **6:** 112-126.

Legendre-Guillemin V, Wasiak S, Hussain NK, Angers A and McPherson PS. **ENTH/ANTH proteins and clathrin-mediated membrane budding.** *Journal of Cell Science* 2004, **117:** 9-18.

Leng Q, Mercier RW, Yao W and Berkowitz GA. **Cloning and first functional characterization of a plant cyclic nucleotide-gated cation channel.** *Plant Physiology* 1999, **121(3):** 753-761.

Leshem Y, Seri L and Levine A. **Induction of phosphatidylinositol 3-kinase-mediated endocytosis by salt stress leads to intracellular production of reactive oxygen species and salt tolerance.** *The Plant Journal* 2007, **51(2):** 185-197.

Letovsky S and Kasif S. **Predicting protein function from protein/protein interaction data: A probabilistic approach.** *Bioinformatics* 2003, **19(1):** i197-i204.

Letunic I, Doerks T and Bork P. **SMART: recent updates, new developments and status in 2015.** *Nucleic Acids Research* 2015, **43(D1):** D257-D260.

Li J, Li M, Liang D, Cui M and Ma F. **Expression patterns and promoter characteristics of the gene encoding *Actinidia deliciosa* L-galactose-1-phosphate phosphatase involved in the response to light and abiotic stresses.** *Molecular Biology Reports* 2013, **40(2):** 1473-1485.

Linder JU. **Class III adenylyl cyclases: molecular mechanisms of catalysis and regulation.** *Cellular and Molecular Life Sciences* 2006, **63(15):** 1736-1751.

Liu SH, Wong ML, Craik CS and Brodsky FM. **Regulation of clathrin assembly and trimerisation defined using recombinant triskelion hubs.** *Cell* 1995, **83(2):** 257-267.

Lohi O, Poussu A, Mao Y, Quiocho F and Lehto VP. **VHS domain – a longshoreman of vesicle lines.** *Federation of the European Biochemical Societies Letters* 2002, **513:** 19-23.

Lomovatskaya LA, Romanenko AS and Filinova NV. **Plant adenylate cyclases.** *Journal of Receptors and Signal Transduction* 2008, **28(6):** 531-542.

Londos C, Cooper DMF, Schlegel W and Rodbell M. **Adenosine analogs inhibit adipocyte adenylate cyclase by a GTP-dependent process: Basis for actions of adenosine and methylxanthines on cyclic AMP production and lipolysis.** *Proceedings of the National Academy of Sciences USA* 1978, **75(11):** 5368-5366.

Lory S, Wolfgang M, Lee V and Smith R. **The multi-talented bacterial adenylate cyclases.** *International Journal of Medical Microbiology* 2004, **293(7-8):** 479-482.

Lucigen Corporation. ***E. cloni*® EXPRESS Chemically Competent Cells. *Product Manual MA019 Rev. C. 2014.*** Wisconsin, USA: Lucigen Corporation.

Ludidi N and Gehring C. **Identification of a novel protein with guanylyl cyclase activity in *Arabidopsis thaliana*.** *The Journal of Biological Chemistry* 2002, **278:** 6490-6494.

Lueking C, Holz C, Lehrach GH, and Cahill D. **A system for dual protein expression in *Pichia pastoris* and *Escherichia coli*.** *Protein Expression and Purification* 2000, **20(3):** 372-378.

Luscombe NM, Greenbaum D and Gerstein M. **What is bioinformatics? An introduction and overview.** *Yearbook of Medical Informatics* 2001, 83-100.

Ma S and Bohnert HJ. **Integration of *Arabidopsis thaliana* stress-related transcript profiles, promoter structures, and cell-specific expression.** *Genome Biology* 2007, **8:** R49.

Magrane Michele and the UniProt Consortium. **UniProt Knowledgebase: a hub of integrated protein data.** *Database* 2011, **bar009**.

Maguire ME. **Hormone-sensitive magnesium transport and magnesium regulation of adenylate cyclase.** *Trends in Pharmacological Sciences* 1984, **5:** 73-77.

Makino T, Skretas G, and Georgiou G. **Strain engineering for improved expression of recombinant proteins in bacteria.** *Microbial Cell Factories* 2011, **10:** 32.

Manfield IW, Jen CH, Pinney JW, Michalopoulus I, Bradford JR, Gilmartin PM and Westhead DR. ***Arabidopsis* Co-expression Tool (ACT): web server for microarray-based gene expression analysis.** *Nucleic Acids Research* 2006, **34(2):** W504-W509.

Mao Y, Chen J, Maynard JA, Zhang B and Quiocho FA. **A novel all helix fold of the AP180 amino-terminal domain for phosphoinositide binding and clathrin assembly in synaptic vesicle endocytosis.** *Cell* 2001, **104(3):** 433-440.

Martinez-Atienza J, Van Ingelgem C, Roef L and Maathuis FJM. **Plant cyclic nucleotide signaling: Facts and fiction.** *Plant Signaling and Behavior* 200, **2(6):** 540-543.

Mazel A, Leshem Y, Tiwari BS and Levine A. **Induction of salt and osmotic stress tolerance by overexpression of an intracellular vesicle trafficking protein AtRab7 (AtRabG3e).** *Plant Physiology* 2004, **134(1):** 118-128.

Meier S and Donaldson L. ***Computational-based analysis to associate the function of plant signaling peptides with distinct biological processes.*** 2012. Berlin, Germany: Springer.

Meier S. **The *Arabidopsis thaliana* wall associated kinase-like protein 10 (AtWAKL10) is a dual domain enzyme with guanylyl cyclase activity and cGMP-dependent kinase activity in vitro and a role in plant defense.** *PLoS ONE* 2010, **5:** e8904.

Mes J and Wichers H. ***Immunomodulation by food for mitigating allergic disease. Bioactive Food as Dietary Interventions for Arthritis and Related Inflammatory Diseases: Bioactive Food in Chronic Disease States.*** 2012. Munich, Germany: Academic Press.

Miroux B and Walker JE. **Over-production of protein in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels.** *Journal of Molecular Biology* 1996, **260(3):** 289-298.

Mittag TW, Guo WB and Kobayashi K. **Bicarboante-activated adenylyl cyclase in fluid-transporting tissues.** *American Journal of Physiology* 1993, **264(6):** F1060-F1064.

Morgan JR, Prasad K, Hao W, Augustine GJ and Lafer EM. **A conserved clathrin assembly motif essential for synaptic vesicle endocytosis.** *The Journal of Neurosciences* 2000, **20(23):** 8667-8676.

Morimoto RI, Kline MP, Bimston DN and Cotto JJ. **The heat-shock response: regulation and function of heat-shock proteins and molecular chaperones.** *Essays in Biochemistry* 1997, **32:** 17-29.

Moutinho A, Hussey PJ, Trewavas AJ and Malho R. **cAMP acts as a second messenger in pollen tube growth and reorientation.** *Proceedings of the National Academy of Science USA* 2001, **98:** 10481-10486.

Mpandeli S, de Villers M, Barnard R, Kgakatsi I and Motsepe M. ***Climate change implications for agriculture in South Africa. Land and water management in Southern Africa: Towards better water use in arid and semi-arid areas.*** 2008. Pretoria, South Africa: The African Institute of South Africa.

Mulaudzi T, Ludidi N, Ruzvidzo O, Morse M, Hendricks N, Iwuoha E and Gehring C. **Identification of a novel *Arabidopsis thaliana* nitric oxide-binding molecule with guanylate cyclase activity in vitro.** *Federation of the European Biochemical Societies Letters* 2011, **585:** 2693-2697.

Murphy D. ***Plants, Biotechnology and Agriculture.*** 2011. Oxfordshire, UK: CABI.

Mutwil M, Obro J, Willats WGT and Persson S. **GeneCAT – novel webtools that combine BLAST and co-expression analyses.** *Nucleic Acids Research* 2008, **36(2):** W320-W326.

Nair SA. **Computational biology and bioinformatics: A gentle overview.** *Communications of the Computer Society of India* 2007, **10:** 1-13.

Nakai K and Horton P. **PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization.** *Trends in Biochemical Science* 1999, **24(1):** 34-36.

NCBI. **National Center for Biotechnology Information**. Maryland, USA: US National Library of Medicine. Accessed at http://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=full_report&list_uids=843139 on 06/01/2015 at 21h03.

Neer EJ. **Interaction of soluble brain adenylate cyclase with manganese.** *The Journal of Biological Chemistry* 1979, **254(6):** 2089-2096.

Newton PCD and Edwards ER. ***Plant breeding for a changing environment. Agro-ecosystems in a changing climate.*** 2007. Florida, USA: CRC Press.

Newton RP, Roef L, Witters E and Van Onckelen H. **Cyclic nucleotides in higher plants: the enduring paradox.** *New Phytology* 1999, **143:** 427-455.

Newton RP. ***Mass spectrometric application in the study of cyclic nucleotides in biochemical signal transduction.*** 2009. Florida, USA: CRC Press.

O'Connor TR, Dyreson C and Wyrick JJ. **Athena: a resource for rapid visualization and systematic analysis of Arabidopsis promoter sequences.** *Bioinformatics* 2005, **21(24):** 4411-4413.

Obayashi T, Kinoshita K, Nakai K, Shibaoka M, Hayashi S, Saeki M, Shibata D, Saito K and Ohta H. **ATTED-II: a database of co-expressed genes and cis elements for**

identifying co-regulated gene groups in Arabidopsis. *Nucleic Acids Research* 2007, **35:** D863-D869.

Okamura N, Tajima Y, Soejima A, Masuda H and Sugita Y. **Sodium bicarbonate in seminal plasma stimulates the motility of mammalian spermatozoa through direct activation of adenylate cyclase.** *The Journal of Biological Chemistry* 1985, **260:** 9699-9705.

Paaventham P, Joseph JS, Seow SV, Vaday S, Robinson H, Chua KY and Kolatkar PR. **A 1.Å structure of Fve, a member of the new fungal immunomodulatory protein family.** *Journal of Molecular Biology* 2003, **332(2):** 461-470.

Paliy O and Gunasekera TS. **Growth of *E. coli* BL21 in minimal media with different gluconeogenic carbon sources and salt contents.** *Applied Microbiology and Biotechnology* 2007, **73:** 1169-1172.

Parker R, Phan T, Baumeister P, Roy B, Cheriyath V, Roy AL and Lee AS. **Identification of TFII-I as the endoplasmic reticulum stress response element binding factor ERSF: its autoregulation by stress and interaction with ATF6.** *Molecular Cell Biology* 2001, **21(9):** 3220-3233.

PastanI and Perlman R. **Cyclic adenosine monophosphate in bacteria.** *Science* 1970, **169(3943):** 339-344.

Pegoraro C, da Rosa Farias D, Mertz LM, dos Santos RS, da Maia LC, Rombaldi CV and de Oliveira AC. **Ethylene response factors gene regulation and expression profiles under different stresses in rice.** *Theoretical and Experimental Plant Physiology* 2013, **25(4):** 261-274.

Perlman RL and Pastan I. **Pleitropic deficiency of carbohydrate utilization in an adenyl cyclase deficient mutant of *Escherichia coli*.** *Biochemical and Biophysical Research Communications* 1969, **37(1):** 151-157.

Pope B and Kent HM. **High efficiency 5 min transformation of *Escherichia coli*.** *Nucleic Acids Research* 1996, **24:** 536-537.

Potter JD, Piascik MT, Wisler PL, Robertson SP and Johnsom CL. **Calcium dependent regulation of brain and cardiac muscle adenylate cyclase.** *Analysis of the New York Academy of Sciences* 1980, **356:** 220-231.

Puertollano R, Randazzo PA, Presley JF, Hartnell LM and Bonifacino JS. **The GGAs promote ARF-dependent recruitment of clathrin to the TGN.** *Cell* 2001, **105(1):** 93-102.

Qi Z, Verma R, Gehring C, Yamaguchi Y, Zhao Y, Ryan CA and Berkowitz GA. **Ca^{2+} signalling by plant *Arabidopsis thaliana* Pep peptides depends on AtPepR1, a receptor with guanylyl cyclase activity, and cGMP-activated Ca^{2+} channels.** *Proceedings of the National Academy of Science USA* 2010, **107(49):** 21193-21198.

Qin X, Coku A, Inoue K and Tian L. **Expression, subcellular localization, and *cis*-regulatory structure of duplicated phytoene synthase genes in melon (*Cucumis melo* L.).** *Plant* 2011, **234(4):** 737-748.

Rahman N, Buck J and Levin LR. **pH sensing via bicarbonate-regulated "soluble" adenylyl cyclase (sAC).** *Frontiers in Physiology* 2013, **4:** 343.

Rai M and Padh H. **Expression systems for production of heterologous proteins.** *Current Science* 2001, **80(9):** 2001.

Rath A, Glibowka M, Nadeau VG, Chen G and Deber CM. **Detergent binding explains anomalous SDS-PAGE migration of membrane proteins.** *Proceedings of the National Academy of Sciences USA* 2009, **106(6):** 1760-1765.

Rehm BHA. **Bioinformatic tools for DNA/protein sequence analysis, functional assignment of genes and protein classification.** *Applied Microbiology and Biotechnology* 2001, **57:** 579-592.

Renak D, Gibalova A, Katarzyna S and Honys D. **A new link between stress response and nucleolar function during pollen development in Arabidopsis mediated by AtREN1 protein.** *Plant, Cell and Environment* 2014, **37(3):** 670-683.

Robichon C, Luo J, Causey TB, Benner JS and Samuelson JC. **Engineering *Escherichia coli* BL21 (DE3) derivative strains to minimize *E. coli* protein contamination after purification by immobilized metal affinity chromatography.** *Applied and Environmental Microbiology* 2011, **77(13):** 4634-4646.

Rogl H, Kosemund K, Kuhlbrandt W and Collinson I. **Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography.** *Federation of the European Biochemical Societies Letters* 1998, **432:** 21-26.

Rosano GL and Ceccarelli EA. **Recombinant protein expression in *Escherichia coli*: advances and challenges.** *Frontiers in Microbiology* 2014, **5:** 172.

Rudolph R and Lilie H. ***In vitro* folding of inclusion body proteins.** *The Federation of American Societies for Experimental Biology Journal* 1996, **10(1):** 49-56.

Ruzvidzo O, Dikobe BT, Kawadza DT, Mabadahanye GH, Chatukuta P and Kwezi L. **Recombinant expression and functional testing of candidate adenylate domains.** *Cyclic Nucleotide Signaling in Plants: Methods and Protocols* 2013, **1016:** 13-25.

Sadler SE and Maller JL. **Inhibition of *Xenopus* oocyte adenylate cyclase by progesterone and 2′,5′-dideoxyadenosine is associated with slowing of guanine nucleotide exchange.** *The Journal of Biological Chemistry* 1983, **258:** 7935-7941.

Sankaranarayanan R, Dock-Bregeon AC, Romby P, Caillet J, Springer M, Rees B, Ehresmann C, Ehresmann B and Moras. **The structure of threonyl-tRNA synthetase-tRNA^{Thr} complex enlightens its repressor activity and reveals an essential zinc ion in the active site.** *Cell* 1999, **97(3):** 371-381.

Saravanan N and Devi T. **A survey on biological databases and applications of datamining.** *Australian Journal of Basic and Applied Sciences* 2012, **6(13):** 175-180.

Schaap P. **Guanylyl cyclases across the tree of life.** *Frontiers in Bioscience* 2005, **10:** 1485-1498.

Schade and Pimentel. **Population crash: Prospects for famine in the 21st century.** *Environmental Development and Sustainability* 2010, **12:** 245-262.

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D and Lohmann JU. **A gene expression map of *Arabidopsis thaliana* development.** *Nature Genetics* 2005, **37:** 501-6.

Schmid SL. **Clathrin-coated vesicle formation and protein sorting: An integrated process.** *Annual Review of Biochemistry* 1997, **66:** 511-548.

Schwacke R, Schneider A, Van Der Graaf E, Fischer K, Catoni E, Desimone M, Frommer WB, Flugge UI and Kunze R. **ARAMEMNON, a novel database for Arabidopsis integral membrane proteins.** *Plant Physiology* 2003, **131(1):** 16-26.

Scott PM, Bilodeau PS, Zhdankina O, Winistorfer SC, Hauglund MJ, Allaman MM, Kearney WR, Robertson AD, Boman AL and Piper RC. **GGA proteins bind ubiquitin to facilitate sorting at the *trans*-Golgi network.** *Nature Cell Biology* 2004, **6:** 252-259.

Seaman MNJ, Burd CG and Emr SD. **Receptor signaling and regulation of endocytic membrane transport.** *Current Opinion in Cell Biology* 1996, **8(4):** 549-556.

Seamon KB. **Forskolin and adenylate cyclase: New opportunities in drug design.** *Annual Reports in Medicinal Chemistry* 1984, **19:** 293-302.

Seki M, Narusaka M, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, Muramatsu M, Hayashizaki Y, Kawai J, Carninci P, Itoh M, Arakawa T, Shibata K, Shinagawa A and Shinozaki K. **Functional annotation of a full-length Arabidopsis cDNA collection.** *Science* 2002, **296(5565):** 141-145.

Shah S and Peterkofsky A. **Characterization and generation of *Escherichia coli* adenylate cyclase deletion mutants.** *Journal of Bacteriology* 1991, **173(10):** 3238-3242.

Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Choi IR, Omura T and Kikuchi S. **Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice.** *Plant Cell Physiology* 2011, **52(2):** 344-360.

Shen J, Snapp EL, Lippincott-Schwartz J and Prywes R. **Stable binding of ATF6 to BiP in the endoplasmic reticulum stress response.** *Molecular Cell Biology* 2005, **25(3):** 921-932.

Shi Y, Mowery RA, Ashley J, Hentz M, Ramirez AJ, Bilgicer B, Slunt-Brown H, Borchelt DR and Shaw BF. **Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding.** *Protein Science* 2012, **21(8):** 1197-1209.

Shin JH, Piao CS, Lim CM and Lee JK. **LEDGF binding to stress response element increases αB-crystallin expression in astrocytes with oxidative stress.** *Neuroscience Letters* 2008, **435(2):** 131-136.

Siewert EA and Kechris KJ. **Prediction of motifs based on a repeated-measures model for integrating cross-species sequence and expression data.** *Statistical Applications in Genetics and Molecular Biology* 2009, **8(1):** 1-34.

Singh KB. **Transcriptional regulation in plants: the importance of combinatorial control.** *Plant Physiology* 1998, **118(4):** 1111-1120.

Singh SM and Panda AK. **Solubilization and refolding of bacterial inclusion body proteins.** *Journal of Biosciences and Bioengineering* 2005, **99(4):** 303-310.

Sinha SC and Sprang SR. **Structures, mechanism, regulation and evolution of class III nucleotidyl cyclases.** *Reviews of Physiology, Biochemistry and Pharmacology* 2007, **157:** 105-140.

Sitaramayya A. ***Signal transduction: Pathways, mechanisms and diseases.*** 2010. New York, USA: Springer Science and Business Media.

Smith C. **Striving for purity: Advances in protein purification.** *Nature Methods* 2005, **2:** 71-77.

Song K, Jang M, Kim SY, Lee G, Lee GJ, Kim DH, Lee Y, Cho W and Hwang I. **An A/ENTH domain-containing protein functions as an adaptor for clathrin-coated vesicles on the growing cell plate in Arabidopsis root cells.** *Plant Physiology* 2012, **159(3):** 1013-1025.

Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A and Tyers M. **BioGRID: A general repository for interaction datasets**. *Nucleic Acids Research* 2006, **1(34):** D535-539.

Steegborn C, Litvin TN, Levin LR, Buck J and Wu H. **Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment.** *Nature Structural and Molecular Biology* 2004, **12:** 32-37.

Steegborn C. **Structure, mechanism, and regulation of soluble adenylyl cyclase – similarities and differences to transmembrane adenylyl cyclases.** *Biochemical and Biophysical Research Communication – Molecular Basis of Disease* 2014, **1842(12):** 2535-2547.

Steer ML and Levitzki A. **The control of adenylate cyclase by calcium in turkey erythrocyte ghosts.** *The Journal of Biological Chemistry* 1975, **250:** 2080-2084.

Stein RJ, Ricachenevsky FK and Fett JP. **Differential regulation of the two rice ferritin genes (*OsFER1* and *OsFER2*).** *Plant Science* 2009, **177(6):** 563-569.

Steinberg SF, Chow YK and Bilezikian JP. **Regulation of rat heart membrane adenylate cyclase by magnesium and manganese.** *The Journal of Pharmacology and Experimental Therapeutics* 1986, **237(3):** 764-772.

Sturman LS. **Characterization of a coronavirus I structural proteins: Effects of preparative conditions on the migration of protein in polyacrylamide gels.** *Virology* 1977, **77(2):** 637-649.

Sun Q, Zybailov B, Majeran W, Friso G, Olinares PD and van Wijk KJ. **PPDB, the Plant Proteomics Database at Cornell.** *Nucleic Acids Research* 2009, **37:** D969-974.

Sutherland EW, Robison AR and Butcher RW. **Some aspects of the biological role of adenosine 3′,5′-monophosphate (Cyclic AMP).** *Circulation* 1968, **3:** 279-306.

Swiezawska B, Jaworski K, Pawelek A, Grzegorzewska W, Szewczuk P and Szmidt-Jaworska A. **Molecular cloning and characterisation of a novel adenylyl cyclase gene, *HpAC1*,**

involved in stress signalling in *Hippeastrum hybridum*. *Plant Physiology and Biochemistry* 2014, **80:** 41-52.

Tebar F, Bohlander SK and Sorkin A. **Clathrin assembly lymphoid myeloid leukemia (CALM) protein: Localisation in endocytic-coated pits, interactions with clathrin, and the impact overexpression on clathrin-mediated traffic.** *Molecular Biology of the Cell* 1999, **10(8):** 2687-2702.

Terpe K. **Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial system.** *Applied Microbiology and Biotechnology* 2006, **72(2):** 211-222.

The Gene Ontology Consortium. **Gene ontology: tool for the unification of biology.** *Nature Genetics* 2000, **25(1):** 25-29.

Thierry-Mieg D and Thierry-Mieg J. **AceView: a comprehensive cDNA-supported gene and transcripts annotation.** *Genome Biology* 2006, **7(Suppl 1):** S12.

Tolia NH and Joshua-Tor L. **Strategies for protein coexpression in *Escherichia coli*.** *Nature Methods* 2006, **3(1):** 55-64.

Turcan S, Vetter DE, Maron JL, Wei X and Slonim DK. ***Mining functionally relevant gene sets for analyzing physiologically novel clinical expression data.*** 2011. Hawaii, USA. Pacific Symposium on Biocomputing.

UN Convention on Biological Diversity. Secretariat of the UN Convention on Biological Diversity. United Nations. 1992. New York, USA.

Usadel B, Obayashi T, Mutwil M, Giorgi FM, Bassel GW, Tanimoto M, Chow A, Steinhauser D, Persson S and Provart NJ. **Co-expression tools for plant biology: Opportunities for hypothesis generation and caveats.** *Plant, Cell and Environment* 2009, **32(12):** 1633-1651.

Verma S, Nizam S and Verma PK. **Biotic and abiotic stress signaling in plants.** *Stress signaling in plants: Genomics and Proteomics Perspective* 2013, **1:** 25.

Walker JM and Rapley R. ***Molecular Biomethods Handbook.*** 1998. New Jersey, USA: Humana Press.

Wang JT, Zaki MJ, Toivonen HTT and Shasha D. ***Introduction to data mining in bioinformatics. Advanced Information and Knowledge Processing.*** 2005. London, UK: Springer.

Wang RS, Pandey S, Li S, Gookin TE, Zhao Z, Albert R and Assmann SM. **Common and unique elements of the ABA-regulated transcriptome of Arabidopsis guard cells.** *BioMedical Central Genomics* 2011, **12:** 216.

Wass MN, Kelley LA and Sternberg MJ. **3D Ligand Site: Predicting ligand-binding sites using similar structures.** *Nucleic Acids Research* 2010, **38:** W469-473.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV and Provart NJ. **An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets.** *PloS ONE* 2007 **2(8):** e7-18.

Wong A and Gehring C. **Computational identification candidate nucleotide cyclases in higher plants.** *Cyclic Nucleotide Signalling in Plants: Methods and Protocols* 2013, **1016:** 195-205.

Xiong L, Schumaker KS, Zhu JK. **Cell signaling during cold, drought, and salt stress.** *The Plant Cell* 2002, **14(Suppl):** S165-S183.

Yadav PR and Tyagi R. ***Crop Biotechnology.*** 2006. New Delhi, India: Discovery Publishing House.

Yan J, Wang B, Jiang Y, Cheng L and Wu T. ***GmFNSII*-controlled soybean flavones metabolism responds to abiotic stresses and regulates plant salt tolerance.** *Plant and Cell Physiology* 2013, **55(1):** 74-86.

Yan SZ, Huang ZH, Andrews RK and Tang WJ. **Conversion of forskolin-insensitive to forskolin-sensitive (mouse-type IX) adenylyl cyclase.** *Molecular Pharmacology* 1998, **53(2):** 182-187.

Yang JK and Epstein W. **Purification and characterization of adenylate cyclase from *Escherichia coli* K12.** *The Journal of Biological Chemistry* 1983, **258:** 3750-3758.

Yanson BR. ***New research on signal transduction.*** 2006. New York, USA: Nova Science Publishers.

Yesilirmak F and Sayers Z. **Heterologous expression of plant gene.** *International Journal of Plant Genomics* 2009, **2009:** 1-16.

Zenser TV. **Inhibition of cholera toxin-stimulated intestinal epithelial cell adenylate cyclase by adenosine analogs.** *Experimental Biology and Medicine* 1976, **152(1):** 126-129.

Zhang B and Horvath S. **A general framework for weighted gene co-expression network analysis.** *Statistical Applications in Genetics and Molecular Biology* 2005, **4(1):** 101-113.

Zhang M, Takano T, Liu S and Zhang X. **Abiotic stress response in yeast and metal-binding ability of a type 2 metallothionein-like protein (PutMT2) from *Puccinellia tenuiflora*.** *Molecular Biology Reports* 2014, **41(9):** 5839-5849.

Zhao J, Zhang J, Zhang W, Wu K, Zheng F, Tian L, Liu X and Duan J. **Expression and functional analysis of the plant-specific histone deacetylase *HDT701* in rice.** *Frontiers in Plant Science* 2014, **5:** 764.

Zhou MY and Gomez-Sanchez CE. **Universal TA cloning.** *Current Issues in Molecular Biology* 2000, **2(1):** 1-7.

Zimmermann P, Hirsch-Hoffmann M, Hennig L and Gruissem W. **GENEVESTIGATOR: Arabidopsis microarray database and analysis toolbox.** *Plant Physiology* 2004, **136(1):** 2621-2632.

Zippin JH, Chen Y, Straub SG, Hess KC, Diaz A, Lee D, Tso P, Holz GG, Sharp GWG, Levin LR and Buck J. **CO_2/HCO_3^- and calcium-regulated soluble adenyly cyclase as a physiological sensor.** *The Journal of Biological Chemistry* 2013, **288:** 33283-33291.

Zouhar J and Sauer M. **Helping hands for budding prospects: ENTH/ANTH/VHS accessory proteins in endocytosis, vacuolar transport and secretion.** *The Plant Cell* 2014, **26(11):** 4232-4244.