

**Molecular profiling and functional characterization
of a novel adenylyl cyclase protein in
two different varieties of soybean**

GH Mabadahanye

 **orcid.org/0000-0002-3409-0302**

Thesis accepted in fulfilment of the requirements for the degree
Doctor of Philosophy in Biology
at the North-West University

Promoter: Prof O Ruzvidzo

Co-promoter: Dr BT Dikobe

Co-promoter: Dr TD Kawadza

Graduation ceremony: May 2023

Student number: 23409444

DECLARATION

I, Grace Humbelani Mabadahanye, declare that the thesis entitled "Molecular profiling and functional characterization of a novel adenylyl cyclase protein in two different varieties of *Glycine max*" 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: Grace Humbelani Mabadahanye

Signature

Date

Supervisor: Prof O Ruzvidzo

Signature

Date

ACKNOWLEDGEMENT

I would like to thank GOD, my heavenly father for giving me life, strength, guidance and wisdom. I am where I am today because of you Lord. There were times when I wanted to give up but you gave me strength to continue with my journey. “To God be the glory”.

I must express my appreciation for my supervisor and mentor Professor O Ruzvidzo, for giving me the opportunity to work under his wing, his efforts, assistance, guidance and patience in his kind supervision. I am honoured to work under his supervision. Your guidance, patience and valuable input are highly appreciated.

Dr Bridgette Tshegofatso Dikobe for her ongoing support, invaluable assistance, guidance and motivation towards the success of this project. This would not have been possible without you.

Dr Dave Kawadza for always offering a helping hand and for all the suggestions you gave me.

To my family Dziedzi and my daughters (Wanga and Yoanda), thank you for your support, patience and for giving me space to do this project and the strength to complete it.

To my siblings, Khodani, Tshilidzi, Vhahangwele, thank you for your support, patience and special thanks to my brothers Daniel and Arron for financially supporting and encouraging me both in good and difficult times.

I would also like to thank all the members of Plant Biotechnology Research Group, your support and assistance throughout my project are highly appreciated.

Thanks, are also due to the North-West University (NWU) Bursary, this work would not have been possible without your financial support.

DEDICATION

This thesis is dedicated to my daughters, Wanga, Yoanda, Onnetulusa and to my late mother.

TABLE OF CONTENTS

DECLARATION.....	I
ACKNOWLEDGEMENT	II
DEDICATION	III
LIST OF TABLES.....	VII
LIST OF FIGURES	VIII
LIST OF ABBREVIATIONS.....	X
KEY TERMS.....	XV
INTRODUCTORY SUMMARY	XIX
CHAPTER 1	1
INTRODUCTION AND LITERATURE REVIEW.....	1
1.1 INTRODUCTION.....	1
1.1.1 Background.....	1
1.1.2. Problem Statement	3
1.1.3 Research Aim	4
1.1.4 Research Objectives	4
1.1.5 Significance of the Research Study.....	5
1.1.6 Chapter Outline	6
1.2 LITERATURE REVIEW	8
1.2.1 Adenylyl Cyclase in Plants	8
1.2.2 cAMP in Higher Plants	9
1.2.3 Adenylylyl Cylases in Legumes.....	10
1.2.4 Glycine max (Soya bean).....	11
1.2.5 Plant Stress	14
CHAPTER 2	22
ISOLATION AND CLONING OF THE GMAC GENE FRAGMENTS	22
ABSTRACT	22
2.1 INTRODUCTION.....	22
2.2 MATERIAL AND METHODS.....	24
2.2.1 Source and Description of the Experimental <i>G. max</i> Varieties	24
2.2.2 Growth of <i>G. max</i> Plants	25
2.3 ISOLATION OF THE <i>G. MAX</i> AC GENE FRAGMENTS	27
2.3.1 Extraction of the Total RNA from Two <i>G. max</i> Varieties	27
2.3.2 Preparation and Amplification of the <i>G. max</i> AC Gene Fragments	28
2.3.4 Agarose Gel Electrophoresis of the Amplified <i>G. max</i> AC Gene Fragments.....	30
2.4 CLONING OF THE AMPLIFIED <i>G. MAX</i> AC GENE FRAGMENTS.....	30
2.4.1 Addition of 3'-Adenines Overhangs	30
2.4.2 Ligation of the <i>G. max</i> AC Gene Fragments into the pTrcHis2-TOPO Vector.....	31
2.4.3. Transformation of <i>E. coli</i> One Shot Competent Cells with the pTrcHis2-TOPO: Gm-AC Constructs	31
2.4.4 Screening for the Positive Clones	32
2.4.5 Extraction of the pTrcHis2-TOPO:GmAC Plasmid Constructs from the Transformed <i>E. coli</i> Cells	32

2.4.6 Analysis of the Positive Clones	33
2.4.7 Agarose Gel Electrophoresis of the Positively Confirmed <i>G. max</i> AC Gene Fragments	35
2.5 RESULTS	35
2.5.1 Generation of the <i>G. max</i> Plants	35
2.5.2 Isolation of the GmAC Gene Fragments from the Two <i>G. max</i> Varieties	36
2.5.3 Cloning of the GmAC Gene Fragments from the Two <i>G. max</i> Varieties	37
2.6 DISCUSSION.....	38
2.7 CONCLUSION	40
2.8 RECOMMENDATION.....	40
CHAPTER 3	41
PARTIAL EXPRESSION AND DETERMINATION OF THE ENDOGENOUS ACTIVITIES OF THE GMAC RECOMBINANT PROTEIN FRACTIONS	41
ABSTRACT	41
3.1 INTRODUCTION.....	42
3.2 MATERIAL AND METHODS.....	44
3.2.1 Transformation of <i>E. coli</i> EXPRESS Chemically Competent BL21 (DE3) pLysS Cells with the GmAC Gene Fusion Expression Constructs	44
3.2.2 Recombinant Expression of the GmAC Protein Fractions	45
3.2.3 Determination of the Endogenous AC Activity of the Recombinant GmAC Protein Fractions.....	46
3.3 RESULTS	47
3.3.1 Partial Expression of the Recombinant GmAC Protein Fractions	47
3.3.2 Determination of the Endogenous AC Activity of the Recombinant GmAC Protein Fractions.....	48
3.4 DISCUSSION.....	49
3.5 CONCLUSION	50
3.6 RECOMMENDATION.....	50
CHAPTER 4	51
PURIFICATION AND DETERMINATION OF THE <i>IN VITRO</i> ACTIVITIES OF THE RECOMBINANT GMAC PROTEIN FRACTIONS	51
ABSTRACT	51
4.1 INTRODUCTION.....	52
4.2 MATERIAL AND METHODS.....	53
4.2.1 Large Scale Expression of the Recombinant GmAC Protein Fractions	53
4.2.2 Determination of the Soluble or Insoluble Statuses of the Expressed Recombinant GmAC Protein Fractions	54
4.2.3 Affinity Purification of the Recombinant GmAC Protein Fractions under Non-native Denaturing Conditions	55
4.2.4 Protein Renaturation	57
4.2.5 Elution of the Refolded Recombinant GmAC Protein Fractions.....	58
4.2.6 Concentration and De-salting of the Recombinant GmAC Protein Fractions	58
4.2.7 Characterization of the AC Activity of the Recombinant GmAC Protein Fractions	59
4.3 RESULTS	61
4.3.1 Determination of the Soluble or Insoluble Statuses of the Expressed Recombinant GmAC Protein Fractions	61
4.3.2 Affinity Purification and Refolding of the Recombinant GmAC Protein Fractions	62
4.3.3 Characterization of the AC Activities of the Recombinant GmAC Protein Fractions.....	63

4.4 DISCUSSION.....	64
4.5 CONCLUSION	66
4.6 RECOMMENDATION.....	66
CHAPTER 5	67
COMPLEMENTATION TESTING OF THE RECOMBINANT GMAC PROTEIN FRACTIONS.....	67
ABSTRACT	67
5.1 INTRODUCTION	68
5.2 MATERIALS AND METHODS.....	70
5.2.1 Isolation and Purification of the pTrcHis2-TOPO:GmAC Expression Constructs from the <i>E. coli</i> EXPRESS BL21 (DE3) PLysS DUOs Cells	70
5.2.2 Preparation of the Competent <i>E. coli</i> CyaA SP850 Cells.....	70
5.2.3 Transformation of the Competent <i>E. coli</i> CyaA SP850 Cells	71
5.2.4 Cell Culturing and Phenotypic Analysis on MacConkey Agar	72
5.3 RESULTS	73
5.4 DISCUSSION.....	74
5.5 CONCLUSION	75
5.6 RECOMMENDATION.....	76
CHAPTER 6	77
EXPLORING THE INVOLVEMENT OF THE GMAC PROTEIN FRACTIONS IN STRESS RESPONSE.....	77
ABSTRACT	77
6.1 INTRODUCTION.....	78
6.2 MATERIAL AND METHODS.....	80
6.2.1 Testing for the Ability of GmAC Protein Fractions to Confer Stress Tolerance to <i>E. coli</i>	80
6.3 RESULTS	81
6.4 DISCUSSION.....	83
6.5 CONCLUSION	86
6.6 RECOMMENDATION.....	86
CHAPTER 7	87
OVERALL DISCUSSION AND RECOMMENDATIONS	87
7.1 OVERALL DISCUSSION	87
7.2 RECOMMENDATIONS	89
REFERENCES	91

LIST OF TABLES

Table 2.1: Components of an RT-PCR reaction mix in a final reaction volume of 50 µl used to amplify the GmAC gene fragments from two <i>G. max</i> varieties.....	29
Table 2.2: The Verso 1-step RT-PCR thermal cycling program used for the amplification of the targeted GmAC gene fragments from two <i>G. max</i> varieties.....	29
Table 2.3: PCR reaction mixture to confirm the successful ligation of the GmAC gene fragment inserts from two <i>G. max</i> varieties into the pTrcHis2-TOPO expression vector.....	34
Table 2.4: Reaction components of a PCR reaction mixture to confirm the correct orientation of the GmAC gene fragment inserts from two <i>G. max</i> varieties in the pTrcHis2-TOPO expression vector.....	34
Table 2.5: Thermal cycling program for the step-by-step assessment profile of the successful cloning of the GmAC gene fragment inserts from two <i>G. max</i> varieties into the pTrcHis2-TOPO expression vector.....	34
Table 4.1: Renaturation conditions of the recombinant GmAC protein fractions from two <i>G. max</i> varieties using the BioLogic Duoflow chromatographic system.....	58
Table 4.2: Molecular characterization of the recombinant GmAC protein fractions from two <i>G. max</i> varieties.....	60

LIST OF FIGURES

Figure 2.1: Nucleotide sequence of the XP_003529590 gene showing the GmAC gene fragment sequence that was isolated, cloned and characterized for each of the two studied varieties of <i>G. max</i> in this work.....	26
Figure 2.2: Generation of <i>G. max</i> plants in potting soil under controlled greenhouse conditions.....	36
Figure 2.3: Isolation of gene fragments from the two <i>G. max</i> varieties.....	37
Figure 2.4: Cloning of the GmAC gene fragments from the two <i>G. max</i> varieties.....	38
Figure 3.1: Partial expression of the GmAC recombinant protein fractions from the two <i>G. max</i> varieties.....	47
Figure 3.2: cAMP generated within the un-induced (CONT) and induced (IPTG) competent <i>E. coli</i> BL21 (DE3) pLysS expression cells harbouring the pTRcHis2-TOPO:GmAC fusion expression constructs.....	48
Figure 4.1: Determination of the soluble/insoluble status of the GmAC protein fractions from the two <i>G. max</i> varieties.....	61
Figure 4.2: Purification of the recombinant GmAC protein fractions from the two <i>G. max</i> varieties under non-native denaturing conditions.....	62
Figure 4.3: Characterization of the AC activity of the recombinant GmAC protein fractions isolated from the NS 5909R (1) or PN 1532A (2) <i>G. max</i> varieties.....	63
Figure 5.1: Testing for the ability of the GmAC protein fractions from the two <i>G. max</i> varieties to rescue <i>cyaA</i> mutation in SP850 cells.....	73
Figure 6.1: Determining the role of GmAC in stress response.....	83

Figure 7.1: Sequence analysis of the AC catalytic center.....	88
--	----

LIST OF ABBREVIATIONS

AC	:	Adenylyl cyclase
ANOVA	:	Analysis of variance
AtCIAP	:	<i>Arabidopsis thaliana</i> clathrin assembly protein
AtLRRAC	:	<i>A. thaliana</i> leucine-rich repeat protein
AtPPR	:	<i>A. thaliana</i> pentatricopeptide repeat protein
ATP	:	Adenosine 5'-triphosphate
cAMP	:	Cyclic 3',5'-adenosine monophosphate
cAMP-CRP	:	Cyclic 3',5'-adenosine monophosphate receptor protein
cDNA	:	Copy DNA or DNA complementary to RNA
CDS	:	Coding DNA sequence
cGMP	:	Cyclic 3',5'-guanine monophosphate
Ca²⁺	:	Calcium ion
CO₂	:	Carbon dioxide
Cl⁻	:	Chloride ion
CNGCs	:	Cyclic nucleotide-gated channels
CNC	:	Cyclic nucleotide cyclase
cNMP	:	Cyclic nucleotide monophosphate
CRP	:	cAMP receptor protein

CDS	:	Coding DNA sequence
cvs	:	Column volumes
cya	:	Adenylate cyclase gene
DNA	:	Deoxyribonucleic acid
DTT	:	Dithiothreitol
EDTA	:	Ethylene di-amine tetra-acetic acid
EIA	:	Enzyme immunoassay
ETI	:	Intracellular receptors encoded by resistance genes
<i>G. max</i>	:	<i>Glycine max</i>
GmAC	:	<i>G. max</i> adenylyl cyclase
GTP	:	Guanosine 5'-triphosphate
GST	:	Glutathione S-transferase
HCO₃⁻	:	Bicarbonate
H₂O₂	:	Hydrogen peroxide
His	:	Histidine
HpAC	:	<i>Hippeastrum hybridum</i> adenylyl cyclase protein
HR	:	Hypersensitive response
HSPs	:	Heat shock proteins
IB	:	Inclusion body

IBMX	:	3-Isobutyl-1-methylxanthine
IDP	:	Intrinsically disordered protein
IDRs	:	Intrinsically disordered regions
IPTG	:	Isopropyl- β -D-thiogalactopyranoside
JA	:	Jasmonic acid
K⁺	:	Potassium ion
KCl	:	Potassium chloride
LB	:	Luria-Bertani
MBP	:	Maltose binding protein
MgCl₂	:	Magnesium chloride
Mg₂SO₄	:	Magnesium sulphate
Mn²	:	Manganese
MOPS	:	3-(N-morpholino) propane sulfonic acid
mRNA	:	Messenger ribonucleic acid
MWCO	:	Molecular weight cut-off point
Na⁺	:	Sodium ion
NaCl	:	Sodium chloride
NADPH	:	Nicotinamide adenine dinucleotide phosphate
NaH₂PO₄	:	Sodium dihydrogen phosphate

NbAC	:	<i>Nicotiana benthamiana</i> adenylyl cyclase protein
NCs	:	Nucleotide cyclases
Ni-NTA	:	Nickel-nitrilotriacetic acid
O₂	:	Oxygen
OD	:	Optical density
PAMPs	:	Pathogen-associated molecular patterns
PAL	:	Phenylalanine ammonia lyase
PCD	:	Programmed cell death
PDE	:	Phosphodiesterase
PGPR	:	Plant growth promoting rhizobacteria
PKA	:	Protein kinase A
PMSF	:	Phenylmethylsulfonyl fluoride
PTI	:	Pattern-triggered immunity
PPR	:	Pentatricopeptide repeat
PR	:	Pathogenesis-related
QTL	:	Quantitative trait loci
RNA	:	Ribonucleic acid
ROS	:	Reactive oxygen species
RT-PCR	:	Reverse transcriptase-polymerase chain reaction

SA	:	Salicylic acid
SDS-PAGE	:	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SCREC	:	School Research and Ethics Committee
SOC	:	Super optimal broth
SUMO	:	Short for small ubiquitin-like modifier
TBE	:	Tris/borate/EDTA
TBS	:	Tris-buffered saline
tmAC	:	Transmembrane adenylyl cyclase
TPR	:	Tetratricopeptide repeat
Tris-HCl	:	Tris hydrochloride
tRNA	:	Transfer ribonucleic acid
UTR	:	Untranslated region
UV	:	Ultraviolet light
VICs	:	Voltage-independent channels
ZmPSiP	:	<i>Zea mays</i> pollen-signaling protein
ZmRPP13-LK3	:	<i>Zea mays</i> putative disease-resistance RPP13-like protein 3

KEY TERMS

Abiotic stress: The negative impact of non-living factors on the living organisms such as plants in a specific environment.

Adenylyl cyclases: Enzymes that catalyse the cyclisation of adenosine 5'-triphosphate (ATP) to cyclic adenosine 3',5'-monophosphate (cAMP), which requires the removal of a pyrophosphate (PPi). They are also known as adenylate cyclases.

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

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

Biotic stress: An unfavorable state that occurs to plants as a result of damage done to them by other living organisms.

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

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

Cyclic adenosine 3',5'-monophosphate (cAMP): A cyclic nucleoside that acts as a second messenger, activating other enzymes within the cell. It is formed when the enzyme adenylyl cyclase is activated by the alpha subunit of a G protein.

Copy DNA: A DNA strand that is complementary to the RNA and synthesised by an RNA-dependent DNA polymerase.

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

DNA sequencing is the process of determining the sequence of nucleotide bases (As, Ts, Cs, and Gs) in a piece of DNA

Drought: A period during which precipitation is below average in a given place, resulting in prolonged shortages in the water supply whether atmospheric, surface water or ground water.

Endogenous: Originating from within an organisms' tissues or cells.

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

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.

Gene: A DNA sequence that can be transcribed into a transcript. In the case of protein coding genes, this transcript can be translated into a protein that is fully functional.

***Glycine max*:** A legume that is commonly known as soybean and is an annual herbaceous plant in the Fabaceae family, cultivated for its edible and highly proteinaceous seed.

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.

Microarray: A platform for quantifying gene expression that assays mRNA molecules based on their hybridisation to probes present on an array, typically a glass slide.

Mitochondrion: A membrane-bound organelle present in the cytoplasm of eukaryotic cells, primarily responsible for the generation of ATP.

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

mRNA: A sub-type of RNA that is generated during transcription, where a single strand DNA is decoded by a DNA-dependent RNA polymerase.

Plasmid DNA: A small circular double-stranded DNA molecule that is distinct from the commonly known chromosomal or nuclear DNA.

Primers: Short nucleic acid sequences that provide a starting point for DNA or RNA synthesis in a conventional PCR system.

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.

Recombinant protein: A protein molecule that is encoded by a gene in which the gene is cloned in a system none other than its host that supports expression of such gene.

Regulatory gene: A gene that controls the expression of other genes.

Reverse transcription-polymerase chain reaction (RT-PCR): A molecular method used to convert a short RNA segment into a DNA product termed copy DNA (cDNA) using an RNA-dependent DNA polymerase enzyme.

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.

Soybean: An annual legume that is cultivated for its proteinaceous seeds, forage and soil improvement. It is scientifically known as *Glycine max*.

TBE buffer: is a buffer solution containing a mixture of Tris base, boric acid and EDTA

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

Transcript: A single-stranded RNA molecule resulting from the transcription of a functional gene.

Variety: A plant line that has been produced in cultivation by selective breeding and is normally propagated vegetatively to maintain a certain trait.

INTRODUCTORY SUMMARY

Adenylyl cyclases (ACs) are a special group of enzymes that catalyze formation of the second messenger molecule, 3',5'-cyclic adenosine monophosphate (cAMP) from 5'-adenosine triphosphate (ATP). In plants, cAMP controls various cellular and physiological processes that are key for survival, growth and development. Some of such processes include control of the cell cycle in tobacco, transport of sodium ions via the voltage-independent channels (VICs) in *Arabidopsis thaliana*, stomatal closure in *Vicia faba*, growth of pollen tubes in *Agapanthus umbellatus*, *Lilium longiflorum* and *Zea mays*, activation of the phenylalanine ammonia lyase (PAL) enzyme in French beans and regulation of the phenylpropanoid pathway in *A. thaliana*. cAMP is also involved in stress response primarily via the cyclic nucleotide-gated channels (CNGCs). Apparently, while a number of ACs have been identified in plants and their cAMP-dependent activities reported, no study to date has attempted to work and/or report on AC activities that are variety-specific for a particular plant species. Therefore, considering the importance of soybean (*Glycine max*) as a source of food, feed, oils, fuels and nitrogen, the present work is an attempt to answer this question. This was achieved by isolating fractions of a GmAC protein, annotated to be an AC in this legume, followed by assessment of the possible variety-specific activities. The stated GmAC fractions were specifically isolated from the NS 5909R variety, known to be drought tolerant and PN 1532A variety, known to be pest resistant. The isolated protein fractions were then cloned (Chapter 2) followed by their comparative assessment for endogenous activity (Chapter 3), *in vitro* activity (Chapter 4), complementation activity (Chapter 5) and functional role in stress response (Chapter 6). Thrillingly, across all these various assessments and evaluations, we found that essentially, no significant differences between the two studied protein fractions existed to report and/or establish a possible variety-specific activity for the GmAC protein. Notably, further analysis of the two studied protein fractions through sequencing and sequence analysis (Chapter 7), showed that the catalytic center of

nearly all ACs, including GmAC, is highly conserved thus explaining why the two studied protein fractions herein could not show any variety-specific activity.

Keywords: Adenylate Cyclase, Cultivars, Cloning, Cyclic Adenosine 3',5'-monophosphate (cAMP), Enzyme Immunoassay, *Glycine max*, Motif, Polymerase Chain Reaction (PCR), Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE), Transformation

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

1.1.1 Background

Food shortage is a major problem facing all African countries because most of them are highly vulnerable to droughts and other harsh environmental conditions that render their farming lands unproductive. This aspect in turn affects the biomass of cultivated crops so that each and every year, there is a decrease in food supplies associated with a rapid increase in the size of populations. This imbalance in the “demand and supply” dynamics has seen an unprecedented increase in food prices yet and moreover, most inhabitants of the African continent do rely on crops as their main source of food and in order to sustain themselves. With the present age and nature of environmental stresses around Africa, crop farming is somewhat becoming increasingly unsustainable, and thus strongly calls for more efficient and cost-effective methods to ensure food security for the current and up-coming human generations. Thus, the main reason why an urgent need to use rational approaches to develop crop plants with increased stress tolerance has been very imminent in the areas of plant genetics, plant physiology, plant biochemistry and plant molecular biology, and a realization that only an integrated and systems-based approach could possibly deliver effective biotechnological solutions (Stuhmer *et al.*, 1989).

Plants are important organisms that need to respond and adapt to environmental conditions for survival. Plants have natural cellular signalling systems involved in detecting and responding to various environmental stimuli for gaining stress resistance. Plants react and

respond to different environmental stresses in various ways. Several studies have elucidated that various mobile plant molecules, such as small proteins, peptides, RNAs, metabolites and second messengers, transmit extracellular stimuli from sensing areas for accurately connecting various organs in plants despite the lack of a nervous system (Takahashi and Shinozaki, 2019).

Plants represent the main source of human foods and livestock feeds and purposes to improve them in many agronomic aspects have primarily focused on plant breeding and production. In the recent times, biotechnological revolution has made it possible for plant breeders to access new sources of genetic variability for the development of superior cultivars. It has now become possible to use additional strategies for crop improvement through the introduction and stable integration of heterologous genes in plant cells that regulate important agronomic characteristics. The genetic manipulation of plants allows their adaptation to different environmental stresses, whether biotic or abiotic. Currently, the production of genetically modified plants occupies a prominent place in both, basic and applied plant research. Plants possess their own alternative mechanisms to defend themselves against invasion by pathogens and such mechanisms include physical barriers, secondary metabolites and antimicrobial proteins. Once established, elicited molecules produced and released by the pathogen induce new defences such as cell wall strengthening, phytoalexin production, synthesis of proteins related to plant defense, among (Ferreira *et al.*, 2006). The identification and application of these mechanisms is one of the most effective systems to rapidly improve crop resistance to diseases (Van Loon *et al.*, 2006).

Glycine max (soya bean) is one of the most important legume crops that provide a good source of oils and quality protein for livestock and people. Apart from being consumable, soybean products have been gaining attention for their other additional attributes such as their anti-cancerous properties (Ko *et al.*, 2013), in pharmaceuticals and the protein-based bio-degradable properties for possible consideration as alternatives in the plastic industry (Asgher *et al.*, 2020). Growth and grain yield of this important cash crop are highly affected by water stress (Brevedan and Egli, 2003; Liu *et al.*, 2003). However, plants respond and adapt to abiotic and biotic stresses or stress factors with an array of biochemical, physiological, and molecular modifications, and the soybean plant is no exception. Since the sequencing of the entire soybean genome in 2010 (Schmutz *et al.*, 2010), the majority of its protein coding genes remain experimentally unconfirmed (Chai *et al.*, 2015). Through study and new literature, this research will make a significant contribution to the field of science as a whole. In terms of application, the clarification of the biological/functional functions of the soybean GmAC protein may be used in transgenics or cisgenics to boost growth and productivity of this crucial agronomic crop, assisting in the resolution of issues with food security in the area and beyond. Due to its excellent quality as a source of oil and protein for both humans and animals, soybean is of utmost economic importance.

1.1.2. Problem Statement

The significance of ACs and their catalytic product, cAMP, in plants has long been established, linking these signalling molecules to various key processes such as control of the cell cycle and stress response in tobacco (Ehsan *et al.*, 1998; Ito *et al.*, 2014), transport of sodium ions via the voltage-independent channels (VICs) in *Arabidopsis thaliana* (Maathuis and Sanders, 2001), stomatal closure in *Vicia faba* (Jin and Wu, 1999), growth of pollen tubes in *Agapanthus umbellatus*, *Lilium longiflorum* and *Zea mays* (Rato *et al.*, 2004;

Moutinho *et al.*, 2001) and activation of the phenylalanine ammonia lyase (PAL) enzyme in French beans (Bolwell, 1992). However, no study to date has attempted to establish a variety-specific functioning of this system in any plant species. Therefore, the major focus of this research project was to determine the possible evolutionary functional pattern of a putative AC protein in *Glycine max* through an assessment and evaluation of its functional trends between two known varieties developed through conventional breeding.

1.1.3 Research Aim

The main aim of this study was to molecularly profile the functional pattern of a known AC protein in two different varieties of *Glycine max* and determine if there exists any variety-specific evolutionary pattern of this protein in this legume.

1.1.4 Research Objectives

The following key objectives were pursued to address the set research question or aim:

1. To generate viable seedlings of two varieties of *Glycine max* that were developed through conventional breeding.
2. To isolate gene fragments from the two varieties, harbouring the annotated and functionally confirmed AC catalytic motif and clone them into suitable viable prokaryotic expression systems.
3. To partially express the cloned gene fragments and compare the inherent endogenous AC activities of the resultant GmAC protein fractions.
4. To purify the expressed recombinant GmAC protein fractions and compare their inherent *in vitro* AC activities.

5. To comparatively assess the abilities of the two expressed GmAC recombinant protein fractions to rescue an AC-deficient, non-lactose fermenting *E. coli* mutant strain.
6. To explore and compare the possible involvement of the two GmAC protein fractions in stress response.
7. To synoptically analyze the findings of the undertaken work in relation to the initially set research question or aim of the study.

1.1.5 Significance of the Research Study

The complete molecular profiling and evolutionary characterization of the novel AC in different varieties of *G. max* will strongly contribute towards a better understanding of the general mechanisms by which plants respond and adapt to various environmental conditions, particularly stress.

1.1.6 Chapter Outline

Chapter 1: Introduction and Literature Review

In this Chapter, the background to the research is highlighted, in which the introduction, literature review, problem statement, aim, specific objectives and significance of the study are elaborated.

Chapter 2: Isolation and Cloning of the GmAC Gene Fragments

This chapter outlines the isolation of gene fragments harbouring the AC catalytic motif from the two studied *G. max* varieties and the subsequent cloning of such isolated gene fragments into stable prokaryotic systems.

Chapter 3: Partial Expression and Determination of the Endogenous Activities of the GmAC Recombinant Protein Fractions

This chapter outlines the expression of the two recombinant proteins from the two studied *G. max* varieties followed by a comparative analysis of their inherent endogenous AC activities.

Chapter 4: Purification and Determination of the *In Vitro* Activities of the Recombinant GmAC Protein Fractions

This chapter details the affinity purification of the expressed recombinant protein fractions from the two studied *G. max* varieties followed by a comparative analysis of their inherent *in vitro* AC activities.

Chapter 5: Complementation Testing of the Recombinant GmAC Protein Fractions

This chapter outlines a comparative analysis of the ability of the recombinant GmAC protein fractions from the two studied *G. max* varieties to rescue a non-lactose fermenting *E. coli* host mutant back into a lactose fermenting wild-type strain.

Chapter 6: Exploring the Involvement of the GmAC Protein Fractions in Stress Response

This chapter comparatively explores and presents the functional involvement of the recombinant GmAC protein fractions from the two studied *G. max* varieties in salt stress and oxidative stress.

Chapter 7: General Discussion

This chapter presents a synoptic analysis of the whole study, relating the obtained findings to the original and initially set research question.

1.1.7 Ethical Clearance

Approval to perform this study was obtained from the School Research and Ethics Committee (SCREC) of the North-West University (Mafikeng Campus), Ethics number: NWU-01263-19-A9.

1.2 Literature Review

1.2.1 Adenylyl Cyclase in Plants

Adenylyl cyclase (AC) is the enzyme that synthesizes 2',3'-cyclic adenosine monophosphate or cAMP from adenosine 5'-triphosphate (ATP). cAMP functions as a "second messenger" to relay extracellular signals to intracellular effectors, particularly protein kinase A. Regulation of intracellular concentrations of cyclic AMP is largely a result in controlling adenylyl cyclase. The presence of cAMP and its role as a signalling molecule in plant cells is not in question; however, knowledge of the purine nucleotide cyclases (NCs) responsible for its synthesis had largely remained abstract and poorly described. This study specifically focused on the functional characterization of a novel AC protein in two different varieties of *Glycine max* developed through conventional breeding.

To date, ten higher plant ACs have been identified; a newly putative disease-resistance RPP13-like protein 3 (ZmRPP13-LK3; Zm00001d045512) which participates in abscisic acid-mediated resistance to heat stress in maize (Yang *et al.*, 2021; Zhao *et al.*, 2021), an *A. thaliana* leucine-rich repeat protein (AtLRRAC1; At3g14460) involved in defense response against hemibiotrophic and biotrophics pathogens (Bianchet *et al.*, 2019). Additionally the *A. thaliana* clathrin assembly protein (AtCIAP; At1g68110) predicted to be involved in actin cytoskeletal re-modelling during endocytic internalization (Chatukuta *et al.*, 2018), two *A. thaliana* K⁺-uptake permease proteins (AtKUP7; At5g09400 and AtKUP5; At4g33530) responsible for K⁺ transport (Al-Younis *et al.*, 2015, 2018), the *Nicotiana benthamiana* adenylyl cyclase protein (NbAC; ACR77530), responsible for tabtoxinine- β -lactam-induced cells deaths during wildfire diseases (Ito *et al.*, 2014). The *Hippeastrum hybridum* adenylyl cyclase protein (HpAC1; ADM83595), which plays a role in stress signalling (Swiezawska *et al.*, 2014). AC was also found in an *A. thaliana* pentatricopeptide

repeat protein (AtPPR-AC; AT1G62590), responsible for pathogen responses and gene expressions (Ruzvidzo *et al.*, 2013), and a *Zea mays* pollen-signaling protein (ZmPSiP; AJ307886), responsible for pollen tube growth and reorientation (Moutinho *et al.*, 2001).

The existence of ACs and their activities in higher plants and other photosynthetic organisms have been previously reported and demonstrated in plants such as *Phaseolus vulgaris*, *Medicago sativa*, *Risinus communis*, and *Pisum sativum* by several physiological and biochemical experiments (Carricarte *et al.*, 1988). In Cyanobacteria, cAMP levels quickly responded to environmental changes, and in *Anabaena cylindrica*, cAMP has been shown to act as a second messenger of light signal transduction (Ohmori *et al.*, 2002). In algae, many AC genes have been cloned, and in fungi, cAMP signalling is also known to regulate tip growth (Moutinho *et al.*, 2001). In rice blast fungus, cAMP is involved in pathogenic growth of the fungus, while in barley powdery mildew, cAMP is involved in the formation of germ tubes (Moutinho *et al.*, 2001).

1.2.2 cAMP in Higher Plants

cAMP has been shown to possess important functions in plant stress response and defense mechanisms (Gottig *et al.*, 2009; Thomas *et al.*, 2013). Apart from being implicated in activation of phytoalexin synthesis in sweet potato (Bianchet *et al.*, 2019), cAMP has been shown to be involved in many processes including cell cycle control in tobacco cells. Many reports have shown the identification of cAMP-dependent processes in plant cells mainly by biochemical methods. Nevertheless, the cAMP signalling pathway has largely remained unknown because of the lack of evidence for its biological role and identification of a plant AC. However, molecular techniques have managed to identify members of cAMP-dependent proteins such as protein kinase A-like kinases and CREBs cAMP response

element-binding proteins; but the diversity of known ACs has not facilitated their identification by homology search. The algal and fungal enzymes are soluble, and another soluble form of AC was identified in rats with a catalytic domain very similar to ACs from cyanobacteria and myxobacteria (Moutinho *et al.*, 2001).

1.2.3 Adenylyl Cylases in Legumes

Legumes are an important aspect of global agriculture because they fix atmospheric nitrogen through symbiotic relationships with microorganisms (Schmutz *et al.*, 2010). Rhizobia, which include legumes and soil bacteria, have co-evolved a facultative nitrogen-fixing symbiosis (Tian *et al.*, 2012). Bacterial penetration via root hair infection threads and, in parallel, organogenesis of nodules that are then invaded by bacteria, are required for the symbiosis to take hold. To keep the mutualistic nature of the contact, nodulation and infection must be kept under control. After the microsymbiont has triggered symbiotic plant development, evidence suggests a passive bacterial function in nodulation and infection (Timmers *et al.*, 2000). A cAMP-signaling regulatory cascade consisting of three receptor-like ACs, a Crp-like regulator, and a target gene of unknown function has been discovered in *Sinorhizobium meliloti*, the *Medicago* symbiont (Tian *et al.*, 2012). Previous research found a somewhat different pattern in the *Sinorhizobium meliloti*–*Medicago* symbiosis, indicating that epidermal infection control requires a long-term molecular conversation between the two partners (Ferrarini *et al.*, 2008; Timmers *et al.*, 2000). According to the findings, a plant signal activates a cAMP regulatory cascade in *Sinorhizobium meliloti*, which controls the degree of epidermal infection by altering the plant's susceptibility to infection (Sorroche *et al.*, 2019).

The cascade is activated specifically by a plant signal during nodule organogenesis (Tian *et al.*, 2012). Cascade inactivation results in a hyper infection phenotype consisting of abortive epidermal infection events uncoupled from nodulation (Gage, 2004). These findings show that, in response to a plant signal, rhizobia play an active role in the control of infection. It has been suggested that rhizobia may modulate the plant's susceptibility to infection (Tian *et al.*, 2012). This regulatory loop likely aims at optimizing legume infection. In legumes, cAMP has been reported to possess an important role in the sequence of biological events that regulate nodule formation and functioning (Gehring and Turek, 2017). Upchurch and Elkan (1978) proposed the involvement of cAMP in the regulation of ammonia assimilation in the rhizobia *Bradyrhizobium japonicum*. Signal exchange is central to the establishment and maintenance of the mutualistic communication between the two partners. Legume signals described so far include flavonoids as elicitors of the bacterial nod gene expression. Other studies have shown that the presence of cAMP in the rhizobia and legume nodules has not been yet found except in the case of *B. japonicum* cells (Laranjo *et al.*, 2004). cAMP may have an important role in the sequence of biological events regulating nodule formation and functioning.

1.2.4 *Glycine max* (Soya bean)

Soybean (*Glycine max*) is a legume that is cultivated worldwide, an economically important crop for livestock feed, human consumption and industrial use as a source for the production of biodiesel. The world soybean production improved by 4.6% annually from 1961 to 2007 and reached average annual production of 217.6 million tons in 2005-2007. World production of soybean is predicted to increase by 2.2% annually to 371.3 million tons by 2030 (Masuda and Goldsmith, 2009). Before the 1980s, the USA produced more than 50% of the world's soybeans, but between 2005 and 2007, that percentage fell to just 37%.

However, throughout the same time span, Brazil and Argentina consistently saw large increases in their shares (Masuda and Goldsmith, 2009). With 53.9 million tons, or 24.8% of global production, Brazil is the second-largest producer. With 41.4 million tons produced and 19.0% of the global total, Argentina comes in third. Ninety-two percent of the world's soybean production is produced by the top five nations: the United States, Brazil, Argentina, China, and India (Masuda and Goldsmith, 2009). The agricultural industry of South Africa includes a small but significant and growing portion devoted to soybean. In South Africa, soy production did not begin on a big scale until the late 1990s, and the area planted with soy has increased quickly since then. A supportive agricultural policy environment and rising yields encourage the commercialization and use of agricultural biotechnologies (Dlamini *et al.*, 2014). Despite having outstanding nutritious value, soybean only make up a relatively modest portion of the diet of the typical South African household. Soybean products for human consumption, with the latter being relatively low, are among the main uses of soybean, for oil, full-fat cake, and low-fat cake. Currently, analyses show that the livestock industry consumes the majority of soybean (Dlamini *et al.*, 2014). Soybean consumption for oil and protein is at 25%, whilst 60% is for animal feed. The poultry industry is by far the largest consumer of soybean derived proteins in South Africa. A minor 15% is used for human consumption in South Africa where most consumers access soybean through over-the-counter products, including soya sources, soups, and other nutritious breakfast foods such as yogurt and flavoured soymilk products (Dlamini *et al.*, 2014). However, soybean yield is detrimentally affected by soil salinity around the world, limiting global agricultural production (Cao *et al.*, 2018). Soybean is one of the main sources of edible vegetable oil and high-protein livestock feed, and it is the most important dicot crop due to the high content of oil and quality protein in its seeds (Aghaei *et al.*, 2009). Soybean inherits the character of the most important legume crop due to its massive enrichment with 40% protein and 20% oil. For long periods of time, its cultivation remained restricted to limited areas. Poor

germination and seed viability are severe problems in soybean cultivation, which are responsible for low production. High and better crop standards are crucial for obtaining high yield. It is critical to use good quality seed with elevated seed germination rates to achieve the required number of plants (Saeed *et al.*, 2017).

G. max is also amongst the important legume crops, as it is an excellent source of dietary protein, carbohydrate, vitamins and minerals for both human food and animal feed (Sakai and Kogiso, 2008). However, the production of this crop is low in arid and semi-arid regions where salinity is the main problem. Soybean is a salt sensitive to moderately salt-tolerant crop. In order to produce salt-tolerant cultivars of this crop, it is imperative to identify genes, whose products confer improved salt tolerance (Aghaei *et al.*, 2009). Soybean is also an excellent source of dietary fiber, phospholipids, isoflavones (e.g., genistein and daidzein), phenolic acids, saponins, and phytic acid, as well as a trypsin inhibitor, and contain unique phytochemical components that improve health (Kim *et al.*, 2021). Its effectiveness in preventing chronic diseases like arteriosclerosis, heart disease, diabetes, and senile dementia, as well as treating cancer and suppressing osteoporosis, has been proven (Kim *et al.*, 2021). Furthermore, soybean seeds has antimutagenic, anticarcinogenic, and antibacterial properties, as well as the ability to impact fibrinolytic activity, manage blood pressure, and improve lipid metabolism (Kim *et al.*, 2021). Soybean is crucial for more than just food and health. Soybean seeds are also utilized in the manufacturing of varnishes, paints, inks, soaps, lubricants, and other items in the industrial sector (Cahoon, 2003). It is also known for its ability to amend soil as a legume. Soybean leaves and stems can be grazed ensiled or dried to make hay. The foliage is very palatable to cattle and has a high nutritive value and good digestibility (Maitra, 2018).

1.2.5 Plant Stress

Abiotic and biotic stresses have a harmful and serious influence on biomass production and crop yield (Agarwal *et al.*, 2006). Abiotic stresses, if not at all, rarely inflict damage independently; plants undergoing thermal stress will often experience drought and salinity stresses simultaneously as is the nature of things. It is important, however, and possible to measure the biological consequences of each stress separately as is performed under controlled laboratory settings. As revealed by such experiments, the degree of damage that plants are vulnerable to is largely dependent on the developmental stage at which the stress is enforced. Drought and high temperature are two major environmental factors that severely limit plant productivity worldwide, frequently causing wide economic losses to agriculture. As global climate change progresses, agricultural production worldwide faces serious threats from frequent extreme weather conditions (Chen *et al.*, 2012). Integrated approaches that improve the efficiency of agricultural water use and development of plant varieties that can alleviate the negative impacts of environmental stresses to maintain yield stability are essential to sustain and increase agricultural production.

Salinity is a common phenomenon. It has been reported that 344 million ha of land are affected by salinity, the ability of crops to tolerate saline environment has gained much attention due to the imperative demands for higher crop productivity and land utilization (Faustino *et al.*, 2000). Much effort has been put towards understanding the mechanisms to deal with salt stress, and they also turn on a wide range of stress-responsive genes to combat the osmotic and ionic stressors that salinity causes. These efforts continue to yield valuable knowledge about plant osmotic stress responses, cellular ion homeostasis and expression of genes with salt tolerance (Faustino *et al.*, 2000). The regulatory network for abiotic stress responses in higher plants have been widely studied in model plants such as

A. thaliana. Some homologous components involved in salt stress responses have subsequently been identified in soybean (Phang *et al.*, 2008).

The growth and development of plants is negatively affected by various biotic and abiotic stresses such as salinity, drought, fungi, bacteria and high temperatures. Drought is one of the major environmental factors that cause crop yield loss (Cutforth *et al.*, 2007). Drought negatively affects soybean production worldwide and threatens food and nutrition security (Kunert *et al.*, 2016). Various physiological, morphological and biochemical changes are induced by drought stress in Soybean. Drought can negatively affect N₂ fixation. Besides this, many other physiomorphological traits can be influenced by drought, such as chlorophyll production plant height, and leaf area index. Understanding how crops react to drought stress will help farmers better utilize irrigation water, which will increase yields even in drought-stressed conditions (Basal and Szabó, 2020).

1.2.5.1 cAMP Involvement in Plant Response to Abiotic Stress

The activation of numerous processes at gene, transcript, and protein levels, interconnected by the activity of signaling messengers, is required to obtain information about plant responses to environmental stimuli (Blanco *et al.*, 2020). The cellular metabolism of stressed plants undergoes a remarkable re-organization, allowing stress adaptation. The initiation of oxidative stress and the stimulation of stress-responsive signaling pathways are early alarm stages of plant abiotic stress response (Kacperska, 2004). Cells encounter new recovering homeostasis after the acclimation period, with the manufacture of stress-protective chemicals, at the expense of cellular energy. In this scenario, cAMP can operate as a stress sensor and/or a modulator of cellular metabolism, mostly but not exclusively, via effecting ion channels and regulating ion fluxes (Blanco *et al.*, 2020).

The enhancement of plant salinity tolerance in *Arabidopsis* is mediated by cAMP, which inhibits Na⁺ inflow by deactivating voltage-independent non-selective channels (Pandey and Sanyal, 2020) . The identification and description of two K⁺-uptake permeases in *A. thaliana* further revealed the crucial relationship between K⁺ flux and cAMP generation. AtKUP5 and AtKUP7 are two different types of AtKUP (Al-Younis *et al.*, 2015). Both permeases serve the same purpose and have a functional AC catalytic domain (Al-Younis *et al.*, 2015). AtKUP7 is a K⁺ transporter found in roots that is functional when K⁺ levels are low (Han *et al.*, 2016). Furthermore, AtKUP7 has been characterized as a proton-coupled carrier with an AC function, but whether cAMP synthesis is dependent on K⁺ fluxes and/or if cAMP may influence K⁺ fluxes is yet unknown. K⁺ flux-dependent cAMP build-up in the cytosol is caused by AtKUP5, which can activate downstream components required for K⁺ homeostasis, such as the CNGCs (Al-Younis *et al.*, 2018). The role of cAMP in abiotic stress response is frequently mediated by the regulation of CNGCs (Jha *et al.*, 2016). Surprisingly, these ion channels favor cross-talk between the signaling of these second messengers due to their overlapping binding domains for cyclic nucleotides and calmodulin (Zhao *et al.*, 2005). The functional analysis of *Arabidopsis* CNGC2 reveals that calmodulin can reverse cAMP activation, implying that the physical contact of Ca²⁺ and calmodulin with CNGCs prevents cyclic nucleotide activation of the channels (Li *et al.*, 2005). As a result, cytosolic cAMP, Ca²⁺ and calmodulin can all work together to control currents through CNGCs (Blanco *et al.*, 2020).

Previously, proteomic studies also supported the role of cAMP in monitoring plant response to temperature, as well as to light. cAMP participates as a signalling molecule to the photosynthetic process of acclimation (Blanco *et al.*, 2020). The analyses revealed that,

after cAMP treatment, the most enriched proteins belonged to the GO categories' Response to Stress, Response to Abiotic Stimulus, Response to Salt and Response to Cold. Moreover, there was an enrichment of the category Photosynthesis and Light Reaction Processes in both the up and down regulated cAMP responsive genes. cAMP involvement in photosynthetic pathways has also been described and previously reported that the interaction between cAMP and enzymes is also involved in the Calvin cycle and photorespiration pathway (Blanco *et al.*, 2020). . This is significant since AC activity in chloroplasts was discovered in *Nicotiana tabacum* using a quantitative method based on mass spectrometric analysis. Additionally, it was demonstrated that light had an impact on the buildup of cAMP in oat seedlings, indicating that cAMP could participate in the phytochrome signaling pathway (Witters *et al.*, 2005).

A role for cAMP in plant response to drought was also proposed in wheat, and cAMP signalling is possibly involved in the effect of ABA on protein synthesis during drought (Blanco *et al.*, 2020). It has been also found that cAMP is involved in response to wounding in *H. hybridum* (Świeżawska *et al.*, 2014). An additional important requirement is a clear identification and functional characterization of cAMP-binding proteins involved in the signalling of this second messenger. Currently, many lines of evidence indicate that, in plants, the conversion of cAMP into Ca^{2+} signals via CNGCs is the main signalling mechanism of this cyclic nucleotide (Qi *et al.*, 2010). However, although indications for *bona fide* PKA are lacking, its presence in plants cannot be excluded. New bioinformatics algorithms and molecular tools may provide opportunities to extend the presently scarce knowledge of cAMP-dependent protein kinases (Swiezawska *et al.*, 2018). Moreover, studies on cAMP-dependent changes in transcriptomes, proteomes and phosphoproteomes, as well as metabolomes, will improve the understanding of cAMP

involvement in plant physiological processes, along with acclimation to adverse environmental conditions.

1.2.5.2 Oxidative Stress

Imbalance between the synthesis and build-up of oxygen reactive species (ROS) in cells and tissues, as well as the biological system's ability to detoxify these reactive products, causes oxidative stress. ROS are normally produced as by-products of oxygen metabolism and can play a variety of physiological roles, including cell signaling. Despite this, environmental stressors like UV, ionizing radiation, pollutants, and heavy metals as well as xenobiotics like antiproliferative drugs all contribute to significantly increased ROS production, which creates the imbalance that results in cell and tissue damage (oxidative stress) (Pizzino *et al.*, 2017). Several antioxidants have been exploited in recent years for their actual or supposed beneficial effects against oxidative stress, such as vitamin E, flavonoids, and polyphenols. Oxidative stress is a common feature, which is associated with various abiotic stress factors, and ROS have an important biological role in sensing and activating acclimation mechanisms (Blanco *et al.*, 2020). The superoxide-generating NADPH oxidase integrates Ca^{2+} and ROS signalling, which in turn may be connected to cyclic nucleotides through CNGCs (Miller *et al.*, 2009). Each messenger mutually enhances the induction of the other during abiotic stress conditions, resulting in the propagation of ROS and Ca^{2+} waves across the plasma membrane to establish the proper acclimation response, to which cAMP could directly or indirectly participate (Blanco *et al.*, 2020).

1.2.5.3 Osmotic Stress in Plants

Osmotic differences in the soil can affect plant growth and development. Most agricultural losses are caused by limiting factors for crop growth and development such as drought

stress. Recently, plant producers have found means to reduce chemical fertilizers by new ways such as increasing crop inoculation with plant growth promoting rhizobacteria (PGPR) (Gouda *et al.*, 2018 and García *et al.*, 2017). The use of these innovative breakthrough technologies can help plants growing under these extreme conditions, some types of stress factors such as drought, by increasing root length, which allows for better access to water (García *et al.*, 2017). Osmotic stress can affect different plant species differently e.g., *Triticum aestivum* L (Ahanger *et al.*, 2017). Ahanger *et al.* (2017) reported that plant species when grown under restricted and normal irrigation conditions with potassium applied in difference doses, they demonstrate increased growth and activity of antioxidant enzymes, both in the experimental and control plant cell lines. The outcome showed that this occurred as a result of the potassium increasing the total phenols and tannins, thus strengthening components of the enzymatic and non-enzymatic antioxidant system (Ahanger *et al.*, 2017). The primary source of dietary protein for vegetarians comes from grain legumes including chickpea, pigeon pea, cowpea, dry pea, lentil, mung bean, urd bean, bean (*Phaseolus vulgaris* L.), broad bean, and grass pea, which are consumed in various forms on a regular basis around the world. Additionally, a large portion of the world's processed vegetable oil comes from grain legumes, including the peanut (*Arachis hypogaea* L.) and soybean (Latef and Ahmad, 2015). While lentil and chickpea are considered to be drought-resistant, pea is known to be drought-sensitive (Latef and Ahmad, 2015).

1.2.5.4 Salinity Stress

Salinity is the main abiotic stress that limits plant growth and productivity of plants in many areas of the world because of an increase on the use of poor quality of water for irrigation and soil salinization. Plant adaptation or tolerance to salinity stress involves complex physiological traits, metabolic pathways, and molecular or gene networks (Gupta and Huang, 2014). Salinity is a major stress limiting factor that increases the demand for food

crops. Osmotic stress is the initial stage of salinity stress that causes different physiological changes, such as interruption of membranes, nutrient imbalance, impairing of the ability to detoxify reactive oxygen species (ROS), differences in the antioxidant enzymes and decreased photosynthetic activity, and decrease in stomatal aperture. Salinity stress is also considered as a hyperionic stress. One of the most harmful effects of salinity stress is the accumulation of Na^+ and Cl^- ions in tissues of plants exposed to soils with high NaCl concentrations (Tavakkoli *et al.*, 2010). Salinity stress includes changes in many physiological and metabolic processes, depending on severity and duration of the stress, and ultimately inhibits crop production. Soil salinity is known to represses plant growth in the form of osmotic stress, which is then followed by ion toxicity (Gupta and Huang, 2014).

1.2.5.5 Signaling Under Abiotic Stress

In order for the plants to survive stress, it is important for them to receive a constant signal of their external environment since they are sessile organisms (Wang *et al.*, 2003). Therefore, plants develop a complex network of signalings, which respond to these external stresses either specifically or with more integrated signalling cascades (Bray, 2000). When a plant faces stress, receptors (G-protein-coupled receptors, receptor like kinase, ion channel or histidine kinase) on the membrane will notice the stress signal, which then allows for the activation of signalling molecules such as inositol phosphates, hormones, ROS, cAMP and Ca^{2+} (Schenk and Snaar-Jagalska, 1999). Once the signal reaches the nucleus, genes that are responsive to stress are activated, ultimately resulting in stress tolerance. Naturally, there are two categories of genes that respond to stress: early genes, which are activated within minutes, and delayed genes, which are activated by the early genes' byproducts (Tuteja, 2007). Signalling molecules, including hormones, proteins, nucleotides, amino acids and growth regulators, amongst others, are critical for plant survival since they

allow for the activation of the antioxidant defense systems against stresses (Dmitriev, 2003). In the situation of stressful conditions, signalling molecules are synthesised and interact with other molecules such as osmolytes as well as phytohormones and other metabolites that accumulate and interact with one another to effect a response (Saed-Moucheshi *et al.*, 2014).

Therefore, the main goal of the present study was to molecularly profile the evolutionary functional pattern of a known AC protein in two different varieties of *Glycine max* and possibly elucidate its possible role with respect to environmental stress response and adaptation mechanisms in this important cultivated legume.

CHAPTER 2

ISOLATION AND CLONING OF THE GmAC GENE FRAGMENTS

Abstract

Adenylate cyclases (ACs) are enzymes responsible for the synthesis of 3',5'-cyclic adenosine monophosphate (cAMP) from 5'-adenosine triphosphate (ATP). cAMP is commonly distributed in prokaryotic and eukaryotic systems, playing fundamental roles in the regulation of diverse cellular processes. In plants, cAMP regulates various processes such as enzyme activities, ion-gated channels and gene expression, however, to this day, no study in plants has attempted to check for existence of any possible variety-specific activity/function of ACs in plants and report it. Therefore, in an attempt to pursue this aspect, we hereby report the isolation of fractions of an AC protein (GmAC) from two varieties of *Glycine max* and their molecular cloning into cloning into a pTRcHis2-TOPO expression vector to generate a pTRcHis2-TOPO:GmAC expression construct. for further experimental characterization.

2.1 Introduction

The questions of whether 2',3'-cyclic adenosine monophosphate (cAMP), a product of adenylate cyclases (ACs) from 5'-adenosine triphosphate (ATP), exists in higher plants and if it does, what its role as a signaling molecule is, are now being resolved not least because modern high-resolution detection methods have enabled sensitive and accurate quantifications of cAMP both *in vitro* and *in vivo*. Naturally, the activities of AC and its breakdown enzyme, phosphodiesterase (PDE), tightly control the levels of cAMP in cells

(Conti and Beavo, 2007; Danchin, 1993). As a result of this building up knowledge and information, the physiological activities of cAMP have now been well defined (Botsford *et al.*, 1992), and genes encoding ACs also identified in most diverse species, plants included. For instance, cAMP has been shown to bind to a protein known as the cAMP receptor protein (CRP) in bacteria.

The advancement of more precise procedures with increased detection capabilities has actually provided a proper key on the occurrence of cAMP in plants, putting to rest any previous doubts about its existence. The use of mass spectrometric analysis, which was able to offer unequivocal evidence of the existence of cAMP and other cyclic nucleotides in higher plants, largely dispelled such skepticism (Richards *et al.*, 2002). The activities of plant ACs have also been demonstrated using tandem mass spectrometry (Pacini *et al.*, 1993; Witters *et al.*, 1996). The isotope dilution method, also known as the cAMP protein-binding assay, has been used in the past to evaluate AC activity in soybean (Yunghans and Morre, 1977), particularly the regulation of nodule formation and function (Terakado *et al.*, 1997).

To this point, cAMP levels have been linked to a variety of physiological processes in plants, according to reports from various laboratories worldwide. Such processes include the upregulation of cytosolic Ca^{2+} concentration in *Nicotiana plumbaginifolia* (Volotovski *et al.*, 1998), as well as the downregulation of Na^{+} influx and accumulation in *Arabidopsis* roots among others (Isner and Maathuis, 2011). The growing identification of cyclic nucleotides has revealed more about their role in the perception of external biotic and abiotic stimuli via signal amplification and transmission to appropriate responses (Gehring and Turek 2017). Many studies have revealed the role of endogenous cAMP in plant defense responses, resulting in increased transcript levels of defense-related genes (Jianget *et al.*, 2005). The

ability to examine and determine the AC activities of novel ACs in plants has greatly been aided by the cloning of AC genes and their production as recombinant proteins in appropriate vectors (Kasahara *et al.*, 2016).

However, regardless of the fact that a number of plant ACs have been discovered, not much is known concerning their molecular dynamics, particularly their evolutionary functions between different varieties of plant species. What is only currently known is that some essential residues of the catalytic centers of these weakly conserved proteins have survived divergent evolution and have been integrated into bigger multi-domain proteins (Zhang *et al.*, 2012). Therefore, in an effort to pursue this aspect, we hereby report the isolation of fractions of an AC protein (GmAC) from two varieties of *Glycine max* followed by their molecular cloning of an AC-containing fragment of the XP_003529590 gene (GmAC) into the pTrHis2-TOPO to generate the pTrHis2-TOPO:GmAC expression construct.

2.2 Material and Methods

2.2.1 Source and Description of the Experimental *G. max* Varieties

The two *G. max* varieties (NS 5909R & PN 1532A) that were used in this study were obtained from the Department of Crop Sciences, School of Agricultural Sciences, North-West University, where they were developed, through conventional breeding, to possess particular traits. Variety NS 5909R was developed to be drought tolerant while variety PN 1532A was developed to be pest resistant.

2.2.2 Growth of *G. max* Plants

2.2.2.1 Surface Sterilization of *G. max* Seeds

Following standard procedures, a total of 5 NS 5909R and 5 PN 1532A seeds were surface sterilized by separately transferring them into two sterile 50 ml falcon tubes filled with 15 ml 70% ethanol. The tubes were vortexed 3 times for 30 seconds, followed by washing in a sterilization buffer (0.1% sodium dodecyl sulphate (SDS) + 5% bleach). Following this treatment, the surface-sterilized seeds were thoroughly washed and rinsed (at least 4 times) with sterile distilled water.

2.2.2.2 Growth and Maintenance of *G. max* Seedlings

After sterilization, all seeds from each variety were sowed in pots containing 3 parts peat-based potting soil and 2 parts vermiculite . The pots were kept in a growth chamber set at 25°C, 300 ppm CO₂, 60% humidity and 16 hrs light/8 hrs dark conditions for 28 days.

2.2.2.3 Designing and Acquisition of Sequence-Specific Primers

During the period when the *G. max* seeds were germinated and their seedlings maintained, some sequence-specific primers for the AC catalytic motif in this legume were designed and acquired. This was achieved by first retrieving the genomic sequence of the XP_003529590 gene from UniprotKB (<https://www.uniprot.org/>) in FASTA (canonical) format (Figure 2.1A). A search on the retrieved sequence indicated presence of this putative AC catalytic motif in the *G. max* genome and specifically, within its XP_003529590 protein, from the 325 to the 337 amino acid sequence of the protein (Figure 2.1B). From this protein sequence, two sequence-specific primers were then manually designed at about 50 base pairs away from the putative motif both sides (Figure 2.1). Both the forward and reverse primers were

carefully designed to ensure that they contained a 5' G instead of any other base and/or also not a 5' T at all, since the subsequent downstream addition of the 3' A-overhangs using *Taq* polymerase during the T-A cloning process is most efficient at adding a non-template 3' A next to a C and not at all when the neighbouring base is an A (Brownstein *et al.*, 1996). The designed primers were then sent to Inqaba Biotechnological Sciences (Pretoria, Republic of South Africa) for chemical synthesis and subsequent supply.



C

Forward Primer
5'-gAA ATA gAA AAg TTg AgA TCC AAT TCA ATg-3'

Reverse Primer
5'-gTg Tgg AAT ATC TTC CTT TCT gTC CAg gTC-3'

Figure 2.1: Nucleotide sequence of the XP_003529590 gene showing the GmAC gene fragment sequence that was isolated, cloned and characterized for each of the two studied varieties of *G. max* in this work. **A.** The bolded red highlighted sequence represents the annotated AC catalytic motif of the XP_003529590 gene, while the bolded green highlighted sequences show the priming sites of the targeted and intended GmAC gene fragments. **B.** A complete amino acid sequence of the GmAC protein showing its annotated AC catalytic centre (red highlight and bold) and the priming sites of this targeted catalytic center (green highlights and bold). **C.**

Nucleotide sequences of both the forward and reverse primers for this targeted catalytic center are also presented (blue highlights).

2.3 Isolation of the *G. max* AC Gene Fragments

2.3.1 Extraction of the Total RNA from Two *G. max* Varieties

The targeted GmAC gene fragments from the two studied *G. max* varieties were synthesized from the total *G. max* mRNA obtained by harvesting plant leaf material from the 28-day-old plants using the Thermo Scientific Gene Jet Plant RNA Purification Mini Kit, and according to the manufacturer's protocol (Catalogue # K0801; Thermo Scientific Inc., Massachusetts, USA). About 100 mg of fresh soybean tissue were snap-frozen in liquid nitrogen and ground thoroughly using a sterile pestle and mortar into a fine powder. The grounded *G. max* tissue powder was then immediately transferred to a 1.5 ml microcentrifuge tube. 500 µL of Plant Lysis solution (supplemented with 10 µl of 2 M DTT) was added onto the ground material in order to release as much RNA as possible from the plant cells and prevent RNA degradation. The mixture was vortexed for 20 seconds vigorously at high speed. The mixed contents were then incubated at 56°C for 3 minutes, then centrifuged for 5 minutes at 20 000xg in an LSE High Speed Microcentrifuge (Corning Inc., Amsterdam, Netherlands). A supernatant of around 550 µl was collected and transferred into a clean 1.5 ml microcentrifuge tube containing 250 µl of 96% ethanol, mixed thoroughly by pipetting. The prepared mixture was then transferred to a purification column inserted in a collection tube and centrifuged for 1 minute at 20 000xg. The flow through solution was discarded while the column and its collection tube were re-assembled as before. An amount of 700 µl of Wash Buffer 1 were added to the purification column and centrifuged for 1 minute at 12 000xg. The flow-through and collection tube were discarded while the purification tube was placed into a clean 2 ml collection tube. A total of 500 µl of Wash Buffer 2 were added to the purification column and

centrifuged for 1 minute at 12 000xg. The flow-through was discarded, while the column and collection tube were re-assembled as before. The washing procedure with Wash Buffer 2 was repeated. After centrifugation, the collection tube containing the flow-through was discarded while the purification column was transferred into RNAase-free 1.5 ml Eppendorf tube. mRNA was eluted by adding 50 µl of nuclease-free water directly onto the column membrane and centrifuging for 1 minute at 12,000g. The purification column was discarded while the eluted mRNA was kept. The concentration of the eluted total mRNA was then quantified using a 2000 Nanodrop Spectrophotometer (ThermoScientific Inc., Massachusetts, USA). This was then followed by electrophoresis on a 1.0% (w/v) agarose gel using a Multi-Sub Mini Electrophoresis System (MS7) (Biocom Ltd, Bridge of Weir, UK) for 40 minutes at 80 V to confirm its purity and integrity. The resolved mRNA was visualized under ultraviolet (UV) illumination using a UV-2000 Trans-illuminator (Bio-Rad Laboratories, California, USA). The collection microfuge tube with the eluted mRNA was finally stored at -80°C for further use.

2.3.2 Preparation and Amplification of the *G. max* AC Gene Fragments

The *G. max* total mRNAs extracted from the leaves were used as templates to generate their respective copy DNAs (cDNAs). The generated cDNAs were then used together with the acquired sequence-specific primers to amplify the targeted GmAC gene fragments in a reverse transcription polymerase chain reaction (RT-PCR) system using the Thermo Scientific Verso 1-Step RT-PCR Reddy Mix RT-PCR Kit and following the manufacturer's instruction (Catalogue # AB-1454/LD/A; Thermo Scientific Inc., California, USA). The used reaction mixtures and cycling conditions are shown below in Tables 2.1 and 2.2 respectively.

Table 2.1: Components of an RT-PCR reaction mix in a final reaction volume of 50 μ l used to amplify the GmAC gene fragments from two *G. max* varieties.

Composition	Volume (μ l)	Final Concentration
Verso Enzyme Mix	1	
1-Step PCR Master Mix	25	1X
Forward Primer (10 μ M)	1	200 nM
Reverse Primer (10 μ M)	1	200 nM
RT Enhancer	2.5	
Water (PCR Grade)	19.5	
Template (RNA)	1	1 ng
Total Volume	50	

Table 2.2: The Verso 1-step RT-PCR thermal cycling program used for the amplification of the targeted GmAC gene fragments from two *G. max* varieties.

Step	Temperature ($^{\circ}$ C)	Time	Cycles
cDNA Synthesis	50	15 minutes	1
Thermo-Start Activation	95	15 minutes	1
Denaturing	95	20 seconds	45
Annealing	50	30 seconds	
Extension	72	1 minute	

Final Extension	72	5 minutes	1
-----------------	----	-----------	---

2.3.4 Agarose Gel Electrophoresis of the Amplified *G. max* AC Gene Fragments

The amplified RT-PCR products from the two *G. max* varieties were separated on a 1% agarose gel supplemented with 0.5 µg/ml ethidium bromide. All samples were resolved against a 100 bp Gene-Ruler DNA ladder (Thermo Scientific Inc., California, USA) while immersed in 1X TBE buffer at 80 volts and 250 mA for 50 minutes. The gel was then visualized under UV light using the UV 2000 Trans-illuminator System (Bio-Rad Laboratories Inc., California, USA) and images finally captured with a ChemiDoc Imaging System (Bio-Rad Laboratories Inc., California, USA).

2.4 Cloning of the Amplified *G. max* AC Gene Fragments

2.4.1 Addition of 3'-Adenines Overhangs

A total volume of 1 µl *Taq* polymerase was added to 40 µl of each individual RT-PCR product reaction mixtures and the mixtures then incubated at 72°C for 10 minutes on a C1000 Thermo-cycler System (Bio-Rad Laboratories Inc., California, USA). By adding single A-overhangs to the GmAC gene fragment's three blunt ends, this technique improved the cloning efficiency of the GmAC gene fragment into the pTRcHis2-TOPO expression vector. After that, the reaction mixture was chilled before being used in the pTRcHis2-TOPO TA cloning system.

2.4.2 Ligation of the *G. max* AC Gene Fragments into the pTrcHis2-TOPO Vector

An aliquot of 4 µl for each of the adenylated PCR products was collected and transferred into a fresh PCR tube before 1 µl of the pTrcHis2-TOPO expression vector (Invitrogen, Carlsbad, USA) was added. The ligation mixture was incubated at room temperature for 5 minutes before its subsequent use for the transformation process of competent *E. coli* propagation cells.

2.4.3. Transformation of *E. coli* One Shot Competent Cells with the pTrcHis2-TOPO: Gm-AC Constructs

After the ligation process, about 2 µl of each of the ligation mixtures (pTrcHis2-TOPO:GmAC) were added into an ice-cold Eppendorf tube containing about 40 µl of the One Shot chemically competent *E. coli* cells (Invitrogen Corp., New York, USA). The reaction mixture was mixed gently and incubated on ice for 30 minutes. The mixture was then heat-shocked on a AccuBlock Digital Dry Bath (Labnet International Inc., New Jersey, USA) at 42°C for 30 seconds before being immediately incubated on ice for 5 minutes. Subsequently, the reaction mixture was supplemented with 250 µl of SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM Mg₂SO₄ and 20 mM glucose). The culture was horizontally shaken in an SI-600 Bench-top Shaking Incubator (Lab Companion GMI Inc., Minnesota, USA) at 200 rpm at 37°C for 30 minutes (this incubation step was carried out to allow cells to produce the enzyme β-lactamase, which later and during the screening process, detoxifies ampicillin. About 20 µL aliquots of the transformed cell cultures were then plated onto two Luria-Bertani (LB) agar plates (1% (w/v) agar, 1% (w/v) tryptone powder, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl supplemented with 50 µg/ml ampicillin and 0.5% (w/v) glucose. The plates were incubated

at 37°C overnight in an Incubat 2001651 Table-top Incubator (JP Selecta SA., Barcelona, Spain). The plates were then visually analysed for growth of colonies in the next morning.

2.4.4 Screening for the Positive Clones

The screening procedure was undertaken to firstly, check if the ligation process was successful and secondly, to determine if such a successful ligation process was in the right expression orientation (correct and acceptable reading frame).

2.4.5 Extraction of the pTrcHis2-TOPO:GmAC Plasmid Constructs from the Transformed *E. coli* Cells

Single colonies from the plated LB plates were used to prepare overnight cultures. A colony of cells was picked and inoculated into 10 ml of double strength yeast-tryptone (2YT) media (0.8% (w/v) tryptone powder, 0.5% (w/v) yeast extract and 0.25% (w/v) NaCl) supplemented with 50 µg/ml ampicillin and 0.5% of glucose in 15 ml falcon tubes. The tubes were incubated overnight at 37°C while shaking at 200 rpm. The next morning, cells were harvested by centrifuging at 6 800xg for 5 minutes at room temperature and the supernatant discarded. Plasmid extraction was performed at room temperature using the Gene Jet Plasmid Miniprep Kit (Thermo Fisher Scientific Inc., California, USA). The pelleted cells were resuspended into 250 µl of the resuspension solution supplemented with RNase by pipetting up and down or until no cell clumps remained. The mixture was then transferred into a sterile Eppendorf tube. This method was followed by the addition of 250 µl of Lysis Solution and mixing thoroughly by inverting the tube 4-6 times or until the solution has become viscous and slightly clear. After that, 350 µl of the Neutralization Solution was added and mixed immediately and thoroughly by inverting the tube 4-6 times. The cells were

pelleted by centrifuging for 5 minutes at 16 300xg and the supernatant transferred into the supplied GeneJET spin column. The spin column fitted into a collection tube was centrifuged for 1 minute and its flow-through discarded while the column re-placed into the same collection tube. About 500 µl of Wash Solution (diluted with ethanol prior to first use) were added to the GeneJET spin column and the column further centrifuged for 30-60 seconds. The flow-through was discarded while the column was placed back into the same collection tube. The washing procedure was repeated twice more before the GeneJET spin column was centrifuged for an extra 1 minute to remove any residual Wash Solution. The washed and dried GeneJET spin column was transferred into a sterile Eppendorf tube, where the plasmid DNA was then eluted by adding 50 µl of Elution Buffer to the center of the GeneJET spin column membrane followed by an incubation of 2 minutes at room temperature and centrifugation at 16 300xg for 2 minutes. The used columns were discarded while the purified plasmid DNA was stored at -20°C for further use.

2.4.6 Analysis of the Positive Clones

Confirmation of positive clones for the two varieties was achieved by conventional PCR following the standard *MyTaq* Mix protocol (Bioline, London, UK) to confirm if the GmAC variety-specific gene inserts had been positively cloned into the pTrcHis2-TOPO vector. The reaction mixtures for the processes are shown in Tables 2.3 and 2.4 while the thermal cycling conditions for both processes are shown in Table 4.5 below.

Table 2.3: PCR reaction mixture to confirm the successful ligation of the GmAC gene fragment inserts from two *G. max* varieties into the pTrcHis2-TOPO expression vector.

Component	Volume (μl)
Template (Construct DNA)	1
Insert Primers (20 μM each)	1
MyTaq™ Red Mix, (2X)	25
Water (sdH ₂ O)	Up to 50

Table 2.4: Reaction components of a PCR reaction mixture to confirm the correct orientation of the GmAC gene fragment inserts from two *G. max* varieties in the pTrcHis2-TOPO expression vector.

Component	Volume (μl)
Template (Construct DNA)	1
Insert and Vector Primers (20 μM each)	1
MyTaq™ Red Mix, (2X)	25
Water (sdH ₂ O)	Up to 50

Table 2.5: Thermal cycling program for the step-by-step assessment profile of the successful cloning of the GmAC gene fragment inserts from two *G. max* varieties into the pTrcHis2-TOPO expression vector.

Step	Temperature (°C)	Time (mins)	Cycles
Initial Denaturation	95	1.00	1
Denaturing	95	0.25	

Annealing	50	0.25	25-35
Extension	72	0.17	

2.4.7 Agarose Gel Electrophoresis of the Positively Confirmed *G. max* AC Gene Fragments

PCR products for the two cloned GmAC gene fragments shown in both Tables 2.3 and 2.4 were resolved on an ethidium bromide-stained 1% agarose gel immersed in a 1X TBE buffer at 80 volts and 250 mA for 50 minutes and the samples resolved against a 100 bp GeneRuler DNA ladder. Visualization was undertaken under UV light using a UV trans-illuminator 2000 system (Bio-Rad Laboratories Inc, California, USA) and capturing of the resultant images by a Chemi Doc Imaging System (Bio-Rad Laboratories Inc., California, USA). Double amplification of the targeted gene fragment inserts in both reaction samples (Table 2.3 and 2.4 respectively) were correspondingly confirming their successful and positive cloning into the pTrcHis2-TOPO expression vector.

2.5 Results

2.5.1 Generation of the *G. max* Plants

The two varieties of *G. max* used in this study were grown in sterile potting soil under greenhouse conditions for 28 days. After 28 days, leaf tissue of these plants was then used to isolate gene fragments harbouring the AC catalytic center.

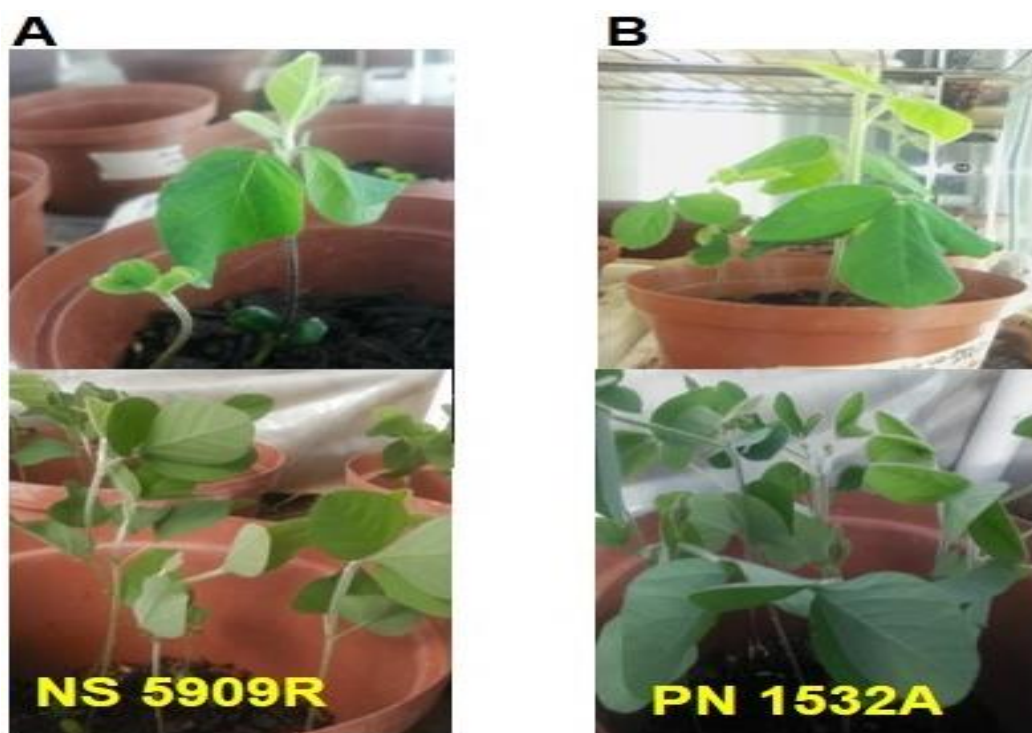


Figure 2.2: Generation of *G. max* plants in potting soil under controlled greenhouse conditions. (A) Variety NS 5909R is known to be drought tolerant and (B) variety PN 1532A known to be pest resistant.

2.5.2 Isolation of the GmAC Gene Fragments from the Two *G. max* Varieties

Gene fragments harbouring the AC catalytic center were specifically isolated through a specialized 1-step RT-PCR system, using the *G. max* total RNA as template. The GmAC gene fragment was then isolated from two types by first extracting total RNA from the plants, and then using the extracted total RNA to create cDNA for the XP 003529590 gene. Using a specialized Verso 1-Step RT-PCR procedure, the produced cDNA and manually created sequence-specific primers were subsequently used to amplify and extract the desired GmAC gene fragment at a 352 bp (Figure 2.3).

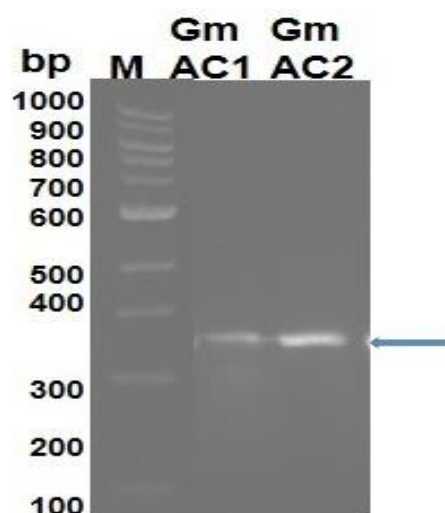


Figure 2.3: Isolation of gene fragments from the two *G. max* varieties. A 1% agarose gel resolution of genes fragments harbouring the AC catalytic center, isolated from the two *G. max* varieties via a specialized RT-PCR system. GmAC1 represents an amplified gene fragment from the NS 5909R variety known to be drought tolerant while GmAC2 represents an amplified gene fragment from the PN 1532A variety known to be pest resistant. M is a 100 bp GeneRuler ladder (Thermo Scientific Inc., Burlington, Canada) while the arrow is marking the expected 352 bp bands of the amplified GmAC1 and GmAc2 genes fragments.

2.5.3 Cloning of the GmAC Gene Fragments from the Two *G. max* Varieties

The isolated *G. max* gene fragments harbouring the AC catalytic centers were cloned into a pTrcHis2-TOPO expression vector, resulting in the production of specific pTrcHis2-TOPO:GmAC fusion expression constructs at a range of 352 bp for both varieties(Figure 2.4).

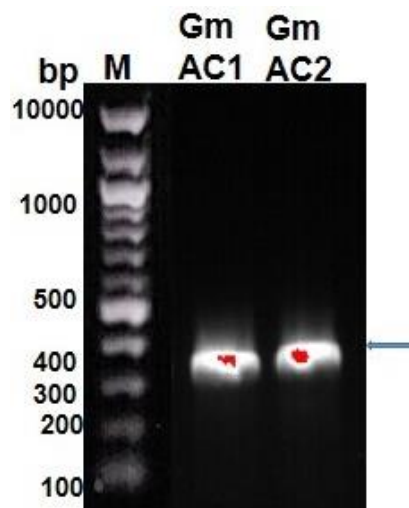


Figure 2.4: Cloning of the GmAC gene fragments from the two *G. max* varieties. Representation of 1% agarose gel used to resolve gene fragments harbouring AC catalytic motifs from two *G. max* varieties after the fragments had been cloned into a pTrcHis2-TOPO vector to generate pTrcHis2-TOPO:GmAC fusion expression constructs. Lane 1 (M) is the 100 bp GeneRuler ladder (Thermo Scientific Inc., Burlington, Canada), lane 2 (GmAC1) is the 352 bp fusion expression construct containing an AC catalytic motif from the NS 5909R variety known to be drought tolerant while lane 3 (GmAC2) is the 352 bp fusion expression construct containing an AC catalytic motif from the PN 1532A variety known to be pest resistant. The arrow marks the two respective variety-specific *G. max* expression constructs amplified and resolved in the analysis.

2.6 Discussion

The generated XP 003529590 cDNA was then used in the same specialized RT-PCR system to amplify and isolate the targeted and intended 352 bp GmAC gene fragment, which was visualized on 1% agarose as is shown in Figure 2.2. Since the main focus of this entire study was on the AC-containing fragment of the XP 003529590 gene from two *G. max* varieties. Direct and extremely effective cloning of these two is made possible by the complementary linearized vector and adenylated PCR (Zhou and Gomez-Sanchez, 2000). However, since both ends of the vector are complimentary to either end of the insert, the TA cloning approach still has the distinct and intrinsic disadvantage of a potential non-directional cloning (Chen and Janes, 2002). In order to clone PCR products, the extracted GmAC gene

fragment was ligated into a pTrcHis2-TOPO plasmid vector using the TA cloning system, which is known as the quickest and most reliable technique. To complement the singly 5'-thymine overhangs on both ends of the expression vector in this method, single 3'-adenine overhangs are often added to the ends of the PCR product first. This was done by placing the sequence GTTCTT on the 5' end of reverse primers, it was noted that it always resulted in adenylation of the 3' end of the forward strand, which then facilitated accurate genotyping and efficient T/A cloning (Brownstein *et al.*, 1996). The resultant plants were then utilized to extract total mRNA, which then was used to isolate the GmAC gene segments of interest, harbouring the targeted AC catalytic center (Figure 2.3). The GmAC gene fragments were then ligated or cloned onto a pTrcHis2-TOPO plasmid vector using the TA cloning approach (Figure 2.4), which has been proved to be the most efficient way of cloning PCR-generated products. After cleavage, a covalent link known as a phosphor-tyrosyl bond is created between the DNA fragment insert and the enzyme. This bond's energy is then employed to push the 5'-hydroxyl end of the split strand to the side. In order to create a recombinant molecule, the Vaccinia topoisomerase 1 can either relinquish the original cleaved bond or religate it to a heterologous acceptor DNA that has a complementary 5'-hydroxyl tail to the adduct (Shuman, 1992).

However, the TA cloning system still has a specific and inherent drawback of a possible non-directional cloning because both ends of the vector are complementary to either end of the insert (Chen *et al.*, 2002). Therefore, to check and confirm if our cloning system was successful (i.e., the correct 5' to 3' directional orientation for proper expression), both GmAC gene fragment (insert) primers (Tables 2.1 and 2.2) and also a single GmAC gene fragment primer in combination with one pTrcHis2-TOPO vector primer (Tables 2.3 and 2.4) for each gene fragment were used to undertake confirmatory PCR tests (Tables 2.5). In this regard, a double amplification of the targeted gene fragment in both reaction systems (Figures 2.3 and 2.4) would automatically confirm its successful cloning into the pTrcHis2-TOPO

expression vector because only correctly oriented inserts can be doubly amplified while wrongly oriented inserts cannot (Invitrogen pTrcHis2 TOPO® Manual, Catalog numbers: K4400-01 and K4400-40).

2.7 Conclusion

The present study successfully isolated gene fragments harbouring the AC catalytic center and cloned them into stable prokaryotic expression systems from two varieties of *G. max*.

2.8 Recommendation

Additional experimental work on the two cloned GmAC gene fragments is required to check and determine if their inherent activities are variety-specific.

CHAPTER 3

PARTIAL EXPRESSION AND DETERMINATION OF THE ENDOGENOUS ACTIVITIES OF THE GmAC RECOMBINANT PROTEIN FRACTIONS

Abstract

Adenylyl cyclases (ACs) and guanylyl cyclases (GCs) generate cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) respectively, which have been recognized as key second messengers affecting many cellular activities in animals for about half a century. This global cyclic nucleotide monophosphate (cNMPs) signaling mechanism has been a source of debate in plants until recently. Histochemical and/or biochemical approaches have historically been used to demonstrate possible adenylate cyclase (AC) activities in higher plants. Attempts to isolate and/or express plant recombinant proteins, followed by their *in vitro* characterization, have recently emerged and got practiced. In *Glycine max*, however, despite the fact that a putative gene (Glyma.07G251000) has been bioinformatically annotated as a possible AC candidate, no study to date has ever attempted to clone it and demonstrate the ability of its resultant protein to produce cAMP from ATP. The chemically competent *E. coli* BL21 (DE3) pLysS cells were then transformed with the resulting expression construct, which was followed by the partial expression of the desired recombinant GmAC protein. The ability of this truncated XP 003529590 (GmAC) protein to produce cAMP from ATP endogenously was then evaluated using enzyme immunoassaying, and it was discovered that the recombinant protein product could actually increase the cellular cAMP levels of those prokaryotic expression hosts by a significant factor of at least 3.0 fold..

3.1 Introduction

There is no doubt that the production of recombinant proteins in microbial systems has revolutionized science in general and biochemistry in particular. Any researcher who embarks on a new project that will need a purified protein, immediately thinks of how to obtain such a protein in a recombinant form. The ability to express and purify the desired recombinant protein in large quantities allows for its biochemical characterization, its use in industrial processes and the development of commercial goods (Rosano and Ceccarelli, 2014).). Theoretically, the steps needed for obtaining a recombinant protein are straightforward. A gene of interest is cloned into whatever expression vector is available, is further transformed into a host of choice, induces expression and consequently, the protein is ready for purification and characterization.

In practice, however, a lot of factors can potentially exert an effect in this whole process. Poor growth of the host, inclusion body (IB) formation, protein inactivity, and including not obtaining any protein at all, are some of the problems often encountered down the pipeline (Rosano and Ceccarelli, 2014). The history of the BL21 and BL21 (DE3) strains has been beautifully documented (Rosano and Ceccarelli, 2014). BL21 cells are deficient in the Lon protease, which reduces the quantity of many foreign proteins. Alongside these two bacterial strains, Luria-Bertani (LB) is the most commonly used medium for culturing *E. coli*. The media is easy to produce since it possesses very rich nutrient contents and its osmolality is optimal for growth at an early log phase. All these features make it suitable for protein production and compensate for the fact that it is not the best option for achieving high cell density cultures. Notably, despite being a rich broth, cell growth tends to stop at a relatively low density. This happens because LB contains scarce amounts of carbohydrates (and other utilizable carbon sources) and divalent cations.

Typically, recombinant protein expression is an important route to determine how protein structures and functions are related. Proteins are usually engineered to be expressed in a suitable expression host such as *E. coli*. Gene expression shows phenotypic changes, which sustain the complexity of organisms (Das and Bansal, 2019). The complexity of gene expression is normally determined by either measuring the number of messenger RNA (mRNA), generally known as the level of expression, or by measuring the number of tissues in which the genes are expressed, normally known as the breadth of expression (Khaitovich *et al.*, 2004). Both approaches are normally capable of revealing the diversity as well as development of specific tissues (Rosano and Ceccarelli, 2014).

In human, chicken and other metazoans, the compactness of genes, coding DNA sequence (CDS) length, gene length, including parameters involved in the intron size, are negatively related to the level of expression and expression breadth (Castillo-Davis *et al.*, 2002). Nevertheless, in plants, precisely *Arabidopsis thaliana* and *Oryza Sativa*, highly expressed genes have longer introns, CDS and untranslated region (UTR) in comparison to genes that are expressed at lower levels (Ren *et al.*, 2006). Furthermore, recombinant proteins can also be identified by Western blotting along expression trials, which is very helpful when the levels of the desired proteins are not high enough to be detected by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

To this day, there is no published literature on molecular cloning, recombinant protein expression, and/or functional characterization of any AC enzyme in *G. max*. The identification and confirmation of the existence and possibly, functional roles of AC enzymes

in agricultural plants, typically requires molecular cloning and recombinant protein expression of AC enzymes from economically important plants like soybean. Confirmation of the biological roles of ACs' enzymatic products, namely the signaling molecule cAMP, can then be employed to improve major features in the production and usage of such important crops.

In this chapter, we describe the partial expression methods of an AC-harboring Glyma.07G251000 gene fragment domain from two varieties of *G. max* in an *E. coli* recombinant system. We then analyze a potential variety-specific AC activity of this recombinant protein in this heterologous prokaryotic system.

3.2 Material and Methods

3.2.1 Transformation of *E. coli* EXPRESS Chemically Competent BL21 (DE3) pLysS Cells with the GmAC Gene Fusion Expression Constructs

After the cloning process (Chapter 2) of the GmAC gene fragments from the two *G. max* varieties studied herein into the pTrcHis2-TOPO expression vector had been practically confirmed, the resultant expression constructs (pTrcHis2-TOPO:GmACs) were then used to chemically transform competent *E. coli* EXPRESS BL21 (DE3) pLysS expression cells. This transformation process was carried out in a sterile Eppendorf tube and in accordance with the manufacturer's protocol and instructions (Lucigen, Wisconsin, USA), whereby 1 µl of the pTrcHis2-TOPO:GmAC gene construct was aseptically added to 100 µl of the ice-cold *E. coli* BL21 (DE3) pLysS cells. The mixture was incubated on ice for 30 minutes before being heat-shocked for 45 seconds at 42°C and immediately placed on ice for 2 minutes. The reaction mixture was supplemented with 250 µl of the room temperature super optimal (SOC) broth and incubated for a 1 hour at 37°C at 225 rpm. After incubation, the

transformation mixture of about 20 µl of the cell culture were plated onto pre-warmed selective LB agar plates [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, 100 µg/ml ampicillin, and 34 µg/ml chloramphenicol, pH: 7.0] and the plates incubated overnight at 37°C in an Incubat 2001651 Table-top Incubator (JP Selecta SA., Barcelona, Spain).

3.2.2 Recombinant Expression of the GmAC Protein Fractions

A single colony of the *E. coli* BL21 (DE3) pLysS DUOs cells harbouring the pTrcHis2-TOPO:GmAC fusion construct from either of the two studied *G. max* varieties was individually 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 37°C overnight at 200 rpm. On the following day, 1 ml of the overnight culture was used to inoculate 20 ml of the selective LB broth. The culture was incubated at 37°C at 200 rpm and up until its optical density (OD₆₀₀) had reached 0.5, as was measured by the Spectronic Helios Epsilon Spectrophotometer (Thermo Electronic Scientific Instruments LLC., Wisconsin, USA). At this desired OD, the culture was immediately chilled on ice and divided into two separate portions of 10 ml each. Protein expression was then induced in one culture by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (SigmaAldrich Corp., Missouri, USA) while the other culture was left un-induced (control). The individually split cultures were then shaken in an incubator for 3 hours at 225 rpm at 37°C. After the 3 hours, 500 µl of each cell culture was separately collected for analysis by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

3.2.3 Determination of the Endogenous AC Activity of the Recombinant GmAC Protein Fractions

Cells from above were analyzed for their inherent levels of cAMP. To inhibit phosphodiesterases, the collected cells (both induced and non-induced) were lysed in 1 ml lysis buffer 1 (Amersham Healthcare, California, USA) supplemented with 2 mM 3- isobutyl-1-methylxanthine (IBMX) (Sigma-Aldrich Corp., Missouri, USA). To speed up the cell lysis process, the samples were agitated for 60 minutes at 200 rpm at room temperature on an orbital shaker. After centrifuging the samples at 9200xg for 5 minutes, 200 µl of the lysates were transferred into new Eppendorf tubes. The lysates were mixed thoroughly with 200 µl of lysis buffer 2 (Amersham Healthcare, California, USA). The reaction mixtures were then divided into 200 µl parts and put into new Eppendorf tubes, where 11 µl of the acetylating reagent (Sigma-Aldrich Corp., Missouri) was added before the mixtures were gently stirred by pulsing. The endogenous cAMP content of the prepared lysates was then determined using a cAMP-linked enzyme immunoassay kit (Catalogue # CA201; Sigma-Aldrich Corp., Missouri, USA), and as stated in the manufacturer's manual. Measurements were taken at 405 nm using a Microplate Reader (Labtech International Limited, East Sussex, UK), and findings analyzed using a one-way analysis of variance (ANOVA).

3.3 Results

3.3.1 Partial Expression of the Recombinant GmAC Protein Fractions

The correctly designed pTrcHis2-TOPO:Gm-AC fusion constructs for the two studied varieties of *G. max* were used to transform chemically competent *E. coli* BL21 (DE3) pLysS DUOs cells, followed by addition of 1 mM IPTG to induce protein expression. The expected 24.79 kDa recombinant GmAC protein fractions were expressed, as illustrated in Figure 3.1 below.

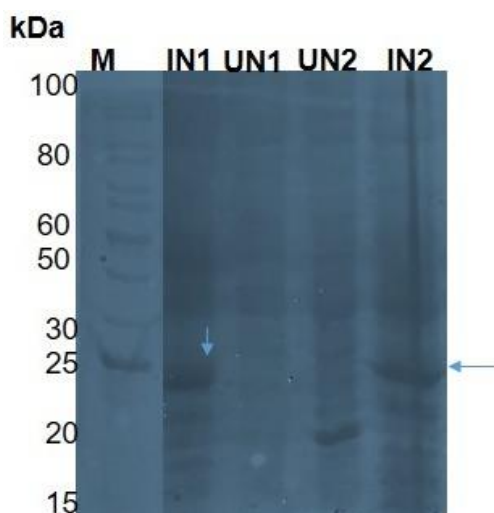


Figure 3.1: Partial expression of the GmAC recombinant protein fractions from the two *G. max* varieties. Induced (IN) and non-induced (UN) *E. coli* BL21 (DE3) pLysS cells harbouring the pTrcHis2-TOPO:GmAC fusion expression constructs were resolved by SDS-PAGE on a 12% gel against a standard molecular weight marker (M). The arrow marks the expressed recombinant GmAC protein fractions from the two studied *G. max* varieties (1 and 2). Variety 1 is NS 5909R known to be drought tolerant while variety 2 is PN 1532A known to be pest resistant.

3.3.2 Determination of the Endogenous AC Activity of the Recombinant GmAC Protein Fractions

Following a successful expression of the desired GmAC protein fractions in transformed *E. coli* EXPRESS BL21 (DE3) pLysS cells, their capacity to promote the conversion of ATP into cAMP within these prokaryotic cells (endogenous) was then tested and determined as is shown in Figure 3.2 below.

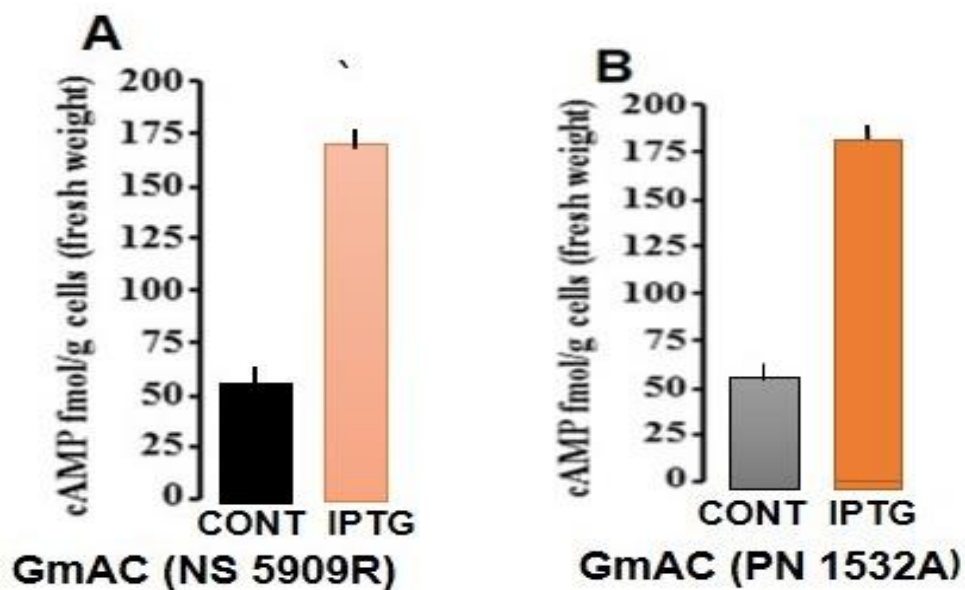


Figure 3.2: cAMP generated within the un-induced (CONT) and induced (IPTG) competent *E. coli* BL21 (DE3) pLysS expression cells harbouring the pTRcHis2-TOPO:GmAC fusion expression constructs. The cells were expressing the GmAC protein fractions from two varieties of *G. max*; NS 5909R known to be drought tolerant and PN 1532A known to be pest resistant. Data are mean values ($n = 3$) as determined by analysis of variance (ANOVA).

3.4 Discussion

Plant tissues have been reported to present numerous challenges for protein extraction because: their rigid cellulosic wall must be sheared off first for the cell contents to be released; the presence of contaminating compounds may cause proteins to clump together; and ultimately, low levels of the targeted protein that degrade, modify or lose activity can be found in highly-complex protein combinations. Because the thick cellulose wall of plant tissues must be sheared off before cell contents can be extracted and released, different protein barriers have been found in plant tissues. Contaminated substances may also cause proteins to deteriorate, change, and degrade (Fido *et al.*, 2004).

Genes that encode the targeted and/or desired plant proteins can also be extracted and used as an alternative. However, the genomic information can be expressed in either bacterial, fungal, or mammalian cell systems. Due to the presence of introns, DNA isolated from eukaryotic organisms for protein production, nevertheless, may result in non-functional protein products. This conundrum can be very troublesome. Only by expressing the desired plant genes in fungi and/or bacteria can this problem be practically overcome. As a result of their eukaryotic nature, animal cells can perform the necessary posttranscriptional changes for targeted recombinant proteins to function (Kusnadi *et al.*, 1997).

In this study, after inducing protein expression with 1 mM IPTG, a recombinant GmAC protein product with a predicted size of 24.79 kDa was produced for both varieties of the *G max* plants studied herein (Figure 3.1). When cells expressing these two versions of the GmAC protein were tested for cAMP production, it emerged that the expressed protein fractions were capable of increasing the levels cAMP in the cells (Figure 3.2) to at least more than 3 folds, presumably through a direct conversion of ATP to cAMP or by activating other resident ACs to generate cAMP.

The current findings are consistent with those of Moutinho *et al.* (2001), Ruzvidzo *et al.* (2013) and Chatukuta *et al.* (2018); who used the pRSET:PSiP, pCRT/7NT-TOPO:AtPPR-AC and TOPO:AtCIAP fusion expression constructs respectively to transform chemically competent BL21 (DE3) and BL21 (DE3) Star pLysS *E. coli* cells and generated almost 3.0-fold levels of endogenous cAMP in all cases. Notably, the levels of cAMP observed herein (Figure 3.2) from the two studied proteins were not significantly different to possibly conclude a probable variety-specific activity.

3.5 Conclusion

In living *E. coli* cells, the isolated, cloned, and partially expressed recombinant GmAC protein from two varieties is capable of generating endogenous cAMP from ATP. This indicates that there is likely another signaling plant molecule that can stimulate the local ACs (*E. coli* ACs) to produce cAMP.

3.6 Recommendation

Since no significant difference in the levels of activity of the two GmAC protein fractions from the two studied *G. max* varieties were detected, further work on these protein fractions is recommended to assess any possible variety-specific activity or function.

CHAPTER 4

PURIFICATION AND DETERMINATION OF THE *IN VITRO* ACTIVITIES OF THE RECOMBINANT GmAC PROTEIN FRACTIONS

Abstract

Despite the fact that cAMP and cGMP were established as significant second messengers in animals about half a century ago, the idea that the cNMP signaling system is ubiquitous and active in plants was a point of contention until recently. These molecules and cAMP, in particular, are now recognized as important signaling molecules in a variety of plant functions, including growth and differentiation, photosynthesis, and biotic and abiotic defense. Historically, histochemical and/or biochemical techniques have been used to assess for possible adenylate cyclase (AC) activities in higher plants. This is because traditional approaches, on the other hand, had been heavily criticized during those days in due to their heavy reliance on intact cells and/or crude cellular extracts, which had always resulted in final findings that were somewhat dubious and/or more akin to guesswork than clear or affirmative conclusions. As a result, it has become clear that pure protein extracts are required in order to give more convincing and trustworthy functional evidence for this crucial group of biological components in plants. In this chapter, to purify the expressed recombinant GmAC protein fractions and compare their inherent *in vitro* AC activities, the His-tagged recombinant GmAC protein fractions, previously cloned and partially expressed in Chapters 2 and 3 respectively, were over-expressed and purified with a nickel affinity gel-based method. Using a cAMP-specific enzyme immunoassaying method, the *in vitro* AC activity of the recombinant GmAC fractions were then determined and ascertained.

4.1 Introduction

In plant–microbe interactions, calcium ion signals have been recently discovered to interact with cGMP (Yuan *et al.*, 2017). Specific response signatures and inter-messenger cooperation emerge from the spatial clustering of stimulus-dependent cyclases and their downstream signalling components, as well as the specific binding of cyclic nucleotides to effector molecules like kinases (Isner and Maathuis, 2016; Kwezi *et al.*, 2011). The need for transient and controlled levels of signalling molecules to generate appropriate responses to environmental and developmental stimuli within defined cellular compartments was highlighted in a study where calcium ion influxes flooded intracellular compartments (Irving *et al.*, 2018; Yuan *et al.*, 2017). When Ca^{2+} is generated in response to environmental stress, it transfers external information quickly from sensor tissues to target tissues (Kudla *et al.*, 2018).

In protein science and studies, recombinant proteins expressed in *E. coli* at small scale levels yield soluble native forms, but those expressed at large scale levels develop aggregates or inclusion bodies (IBs), which are the result of mis-folding of the recombinant proteins in foreign hosts (Zhang *et al.*, 1998). IBs are bio-macromolecule aggregates, usually proteins that can be found in the nucleus, cytoplasm, or periplasm. These protein types are mostly produced by foreign or mutant genes that lack normal post-translational and/or folding modifications. Recombinant proteins are becoming increasingly essential in industrial applications, with *Escherichia coli* serving as the most common bacterial host. IBs, on the other hand, are a common stumbling block in the production of soluble and functional recombinant proteins. Environmental conditions, such as culture temperature, pH and available O_2 , also affect IBs formation during recombinant proteins production in *E. coli* (Bhatwa *et al.*, 2021).

However, interventions such as expression in the periplasmic space of the cell or with a tag, and/or use of other hosts may help the situation. In addition, the recombinant protein may also be labeled before and during cloning to increase its solubility and purification operations. Tags are attached to the protein's N- or C-terminal end and these include the His6, GST, MBP, NusA, Trx, and SUMO tags (Gupta and Shukla, 2016). His-tag, because it comprises 6 or more consecutive negatively charged histidine residues, it is the most often utilized tag due to its efficient affinity purification capabilities on suitable platforms when fused to the protein (Gupta and Shukla, 2016). More so, even in the presence of denaturants like urea and guanidine hydrochloride, the His-tag still enables for a one-step purification system of recombinant proteins (Kneusel *et al.*, 1998). Alongside this, the nickel nitrilotriacetic acid (Ni-NTA) system is the most suitable platform used to purify proteins that have been tagged with the His-tag because it contains an equally positively charged system that is compatible with the His-tag fused onto proteins (Campos *et al.*, 2008). In this study, the Ni-NTA affinity purification approach was therefore, used to purify the recombinant GmAC protein fractions from the two *G. max* varieties that were studied herein, followed by determination and characterization of their AC activities *in vitro*.

4.2 Material and Methods

4.2.1 Large Scale Expression of the Recombinant GmAC Protein Fractions

For large scale expression of the recombinant GmAC protein fractions from either of the two studied *G. max* varieties, 200 µl of cells from the glycerol stocks generated in Chapter 2 were used to inoculate 20 mL of the LB broth media supplemented with 0.4% glucose, 100 µg/ml of ampicillin and 34 µg/ml chloramphenicol. The inoculated culture was then grown

overnight at 37°C in a Labcon shaking incubator at 225 rpm. The next morning and in a 500 mL Erlenmeyer flask, 200 ml of the LB broth media containing 100 µg/ml ampicillin, 34 µg/mL chloramphenicol and 0.4% glucose were inoculated with 2 ml of the overnight culture and incubated at 37°C with agitation at 225 rpm. At an OD₆₀₀ of 0.6, the culture was induced 1 mM IPTG (Sigma-Aldrich Corp., Missouri) and then allowed to grow for a further 2 hours. After the 2 hours, the culture was centrifuged at 16,300xg for 10 minutes, supernatant discarded while the pellet was stored at -20°C for further processing.

4.2.2 Determination of the Soluble or Insoluble Statuses of the Expressed Recombinant GmAC Protein Fractions

The soluble/insoluble state of each of the two expressed recombinant GmAC protein fractions was assessed by lysing the pellet and analyzing the cellular protein components on SDS-PAGE under native non-denaturing conditions. Briefly, the whole pellet of cells produced in section 4.2.1 was thoroughly resuspended in 2.5 ml of sterile Tris-buffered saline (TBS) buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.5), supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 g/ml lysozyme (Catalogue number 62970; Sigma-Aldrich Inc., Missouri, USA). The resulting mixture was then placed on ice for 60 minutes while being vortexed for 10 minutes at medium speed with 1-minute intervals of ice incubation. The capability to slowly rupture cells and releasing their contents was solely dependent on this process. To separate the pellet from the supernatant, the mixture was centrifuged for 2 minutes at 9 200xg. The soluble portion of the protein was maintained in the supernatant, whereas the insoluble component of the protein was kept in the pellet. SDS-PAGE was then used to evaluate the presence and relative amounts of the expressed recombinant GmAC protein in 20 µl of each of the two protein fractions. The presence of the recombinant protein in the supernatant fraction would indicate that the protein is soluble,

whereas its presence in the pellet fraction, would alternately indicate that the protein is insoluble.

4.2.3 Affinity Purification of the Recombinant GmAC Protein Fractions under Non-native Denaturing Conditions

After realizing that the bulk of the GmAC protein (analyzed in section 4.2.2) was in the pellet fraction, the insoluble recombinant GmAC protein complement was then purified under non-native denaturing conditions using a HIS-Select Ni-NTA affinity matrix and in accordance with the manufacturer's instructions (Catalogue # P6611; Sigma-Aldrich Inc., Missouri, USA). Briefly, the pellet was resuspended in 5 mL of Lysis Buffer consisting of 8 M urea, 100 mM NaH₂PO₄, 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 volume to 1 g of the pellet weight. The resuspended cell pellet component was then solubilized by vigorous vortexing for 1 hour at full speed into a crude lysate. The crude lysate was then centrifuged for 30 minutes at 10,000xg at room temperature to yield a cleared lysate through its separation from the overall cell debris. The yielded cleared lysate was then collected and stored at room temperature for further downstream processing.

4.2.3.1 Equilibration of the His-Select Ni-NTA Affinity Matrix

An aliquot of about 2 ml of the 50% (v/v) HIS-Select Ni-NTA affinity bead solution was collected in a sterile 15 ml falcon tube and allowed to settle down at room temperature for about 10 minutes. The beads were then washed three times with 2 ml sterile distilled water to remove all the residual storage ethanol. The bead solution was then equilibrated with 2 mL of Lysis Buffer by mixing the two and allowing the mixture to settle down at room

temperature for 5 minutes, followed by a complete removal and discarding of the buffer, keeping the equilibrated beads. Binding of the recombinant GmAC protein onto the equilibrated HIS-Select Ni-NTA affinity matrix then followed, whereby 5 ml of the cleared lysate (generated previously in section 4.2.3 above) were added onto 2 ml of the equilibrated beads and allowed to mix gently mixed for 1 hour at room temperature on an adjustable rotating revolver (Labnet International Inc., New Jersey, USA) at 30 rpm. During this phase, the recombinant GmAC protein was then bound onto the HIS-Select Ni-NTA affinity beads while the rest of the other bacterial proteins were left un-bound. The mixture was then centrifuged for 10 minutes at 5,000xg and 50 µl of the supernatant (the flow-through) collected for later analysis by SDS-PAGE while the rest of it (flow-through) was discarded.

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

The bound beads were then washed three times to remove all the unbound bacterial proteins after the recombinant GmAC protein was bound onto the equilibrated HIS Select Ni-NTA affinity matrix. Briefly, about 5 ml of Wash Buffer (8 M urea, 100 mM NaH₂PO₄, 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 mixed on a bench revolver at 30 rpm for 5 minutes. The suspension was then centrifuged for 1 minute at 5,000xg before the supernatant (wash 1) was removed and its aliquot (20 µl) collected for analysis by SDS-PAGE. The washing step was repeated two more times (wash 2 and 3) and 20 µl aliquot of the washes collected and saved for analysis by SDS-PAGE. Additionally, a 20 µl aliquots of the washed beads was also saved for analysis by SDS-PAGE.

4.2.4 Protein Renaturation

The recombinant GmAC protein fractions were denatured as a result of the urea pellet solubilization described in section 4.2.2. As a result, a refolding or renaturation procedure was required to restore their functional states.

4.2.4.1 Renaturation of the Recombinant GmAC Protein Fractions

The recombinant GmAC protein-carrying HisPur Ni-NTA resins (section 4.2.3.2) were first resuspended in 10 column volumes (cv) of Refolding Buffer A (8 M Urea, 200 mM NaCl, 50 mM Tris-HCl, 500 mM glucose, 20 mM β -mercaptoethanol, pH: 8.0). After that, the resuspended slurry was placed into a clean XK16 column (Bio-Rad Laboratories Inc, California, USA) and manually attached to a Biologic DuoFlow Chromatography System (Bio-Rad Laboratories Inc., California, USA). The parameters listed in Table 4.1 were then used to produce a linearized gradient system. During this process, gradient Buffer A was diluted to 0 M urea concentration with a Refolding Buffer B (200 mM NaCl, 50 mM 79 Tris-HCl, 500 mM glucose, 0.05 percent (w/v) polyethyglycol, 4 mM reduced glutathione, 0.04 mM oxidized glutathione, and 0.5 mM phenylmethylsulphonyl fluoride), to promote a regulated renaturation process of the denatured recombinant GmAC protein. Both buffer systems were coupled to a controllable BioLogic Duo-Flow Chromatographic System. As a result, the purified recombinant GmAC protein was transformed to its soluble and biologically active state during this procedure.

Table 4.1: Renaturation conditions of the recombinant GmAC protein fractions from two *G. max* varieties using the BioLogic Duoflow chromatographic 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	0.00%
Target concentration of Buffer A	100%
Length of gradient	900 min

4.2.5 Elution of the Refolded Recombinant GmAC Protein Fractions

The bound beads in the column were first washed with 10 column volumes (cvs) of Washing Buffer (200 mM NaCl, 50 mM Tris-HCl; pH: 8.0, 20% (v/v) glycerol, 0.5 mM PMSF) in order to remove any residual refolding agents from the matrix. The washed and refolded recombinant GmAC 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) utilizing an elution process that was directly coupled to the Duo-Flow system.

4.2.6 Concentration and De-salting of the Recombinant GmAC Protein Fractions

About 2 mL of the eluted recombinant GmAC protein were added into a Spin-XUF concentrating and de-salting device with a molecular weight cutoff point (MWCO) of 3.0 kDa (Corning Corp., New York, USA), followed by centrifugation at 2,540xg (Hermle Labortechnik, Wehingen, Germany) at 4°C and up until the final volume was down to 200 µl. Protein concentration was determined using a 2000 Nanodrop Spectrophotometer. After

that, the concentrated and de-salted protein fraction was transferred into a new sterile 1.5 ml microfuge tube (Thermo Scientific Inc., Massachusetts, USA). A small portion of the eluted protein fraction (20 μ l) was then put aside for SDS-PAGE analysis, while the rest was kept at -20°C for later use.

4.2.7 Characterization of the AC Activity of the Recombinant GmAC Protein Fractions

The cAMP-linked enzyme immunoassaying (EIA) system was used to characterize the AC activity of purified and refolded recombinant GmAC protein *in vitro*, following the manufacturer's instructions and based on its acetylation process (Catalogue #CA201: Sigma-Aldrich Inc., Missouri, USA). Molecular components such as GTP, Mg^{2+} , Mn^{2+} , Ca^{2+} , HCO_3^- and F^- were used to assess and determine the functional effects of various ionic and/or molecular components on the enzymatic activity of the recombinant GmAC protein.

4.2.7.1 Preparation of Samples and Enzyme Immunoassaying

In sterile 1.5 ml microfuge tubes, 200 μ l reaction mixtures were made as detailed in Table 4.2 below. Each tube included 50 mM Tris-HCl (pH: 8.0) for buffering and 2 mM IBMX to prevent phosphodiesterases from degrading cAMP, 10 μ g of the recombinant protein, 5 mM $MgCl_2$ or $MnCl_2$ and 1 mM ATP and/or GTP. After a 20-minute incubation period at room temperature, the reaction system was stopped by adding 1 mM EDTA to chelate away all divalent metal ions and therefore, remove the co-factors required for the recombinant GmAC's enzymatic action. To inactivate the protein, the samples were also boiled for 5 minutes. To clarify the boiled samples, the sample mixtures were centrifuged for 5 minutes at 16,300 x g. An acetylating reagent (2:1), triethylamine:acetic anhydride (v/v), was then

added to the samples at a volumetric ratio of 1:20 acetylating reagent: sample, followed by vortexing for 2 seconds at high speed. The resultant solution was then tested for cAMP concentration using the cAMP-linked enzyme immunoassaying kit (Catalog # CA201). Readings were taken in triplicates at 405 nm using a microplate reader (Labtech, International Limited, East Sussex, UK).

Table 4.2: Molecular characterization of the recombinant GmAC protein fractions from two *G. max* varieties.

	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
MnCl ₂	5mM	5mM	-	5mM	5mM	5mM	5mM	5mM	5mM
Protein	-	25µg	25µg	25µg	25µg	25µg	25µg	25µg	25µg
ATP	1mM	1mM	1mM	1mM	1mM	1mM	-	1mM	1mM
MgCl ₂	-	-	5mM	-	-	-	-	-	-
CaCl ₂	-	-	-	250µM	-	-		-	-
GTP	-	-	-	-	-	1mM	1mM	-	-
NaHCO ₃	-	-	-	-	50mM	-	-	-	-
NaF	-	-	-	-	-	-	-	-	10mM
AlCl ₃									+30µM

4.3 Results

4.3.1 Determination of the Soluble or Insoluble Statuses of the Expressed Recombinant GmAC Protein Fractions

The recombinant GmAC protein fractions were mostly over-expressed as insoluble inclusion bodies in the BL21 (DE3) pLysS *E. coli* cells, resulting in some aggregated and non-functional protein products. This was determined by the solubility/insolubility test described in section 4.2.2, in which the bulk of the expressed recombinant GmAC protein components could only be found in the insoluble cellular fraction after cell lysis and fractionation in a native buffer. An SDS-PAGE of the soluble and insoluble cellular contents is shown in Figure 4.1 below.

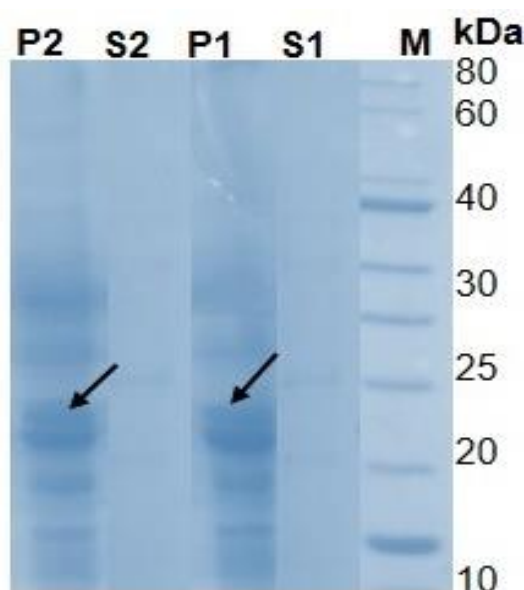


Figure 4.1: Determination of the soluble/insoluble status of the GmAC protein fractions from the two *G. max* varieties. SDS-PAGE of the soluble (S) and insoluble (P) components of cells expressing the recombinant GmAC protein fractions. The analyzed GmAC protein fractions were from the NS 5909R (1) or PN 1532A (2) *G. max* varieties known to be drought tolerant and pest resistant respectively. M is the standard molecular weight marker and arrows mark the resolved GmAC protein bands.

4.3.2 Affinity Purification and Refolding of the Recombinant GmAC Protein Fractions

The poly-histidine C-tagged fusion products (GmAC protein fractions) were purified (Figure 4.2, below) under non-native denaturing conditions using an HIS-Select Ni-NTA affinity matrix system and based on the manufacturer's protocol and instructions (Catalogue # P6611; SigmaAldrich Inc., Missouri, USA). This decision was undertaken after determining that the recombinant GmAC protein fractions were wholly expressed in form of inclusion bodies (Figure 4.1). The purified protein fractions were then refolded into their native soluble forms and later eluted off the Ni-NTA matrix as is shown below in Figure 4.2.

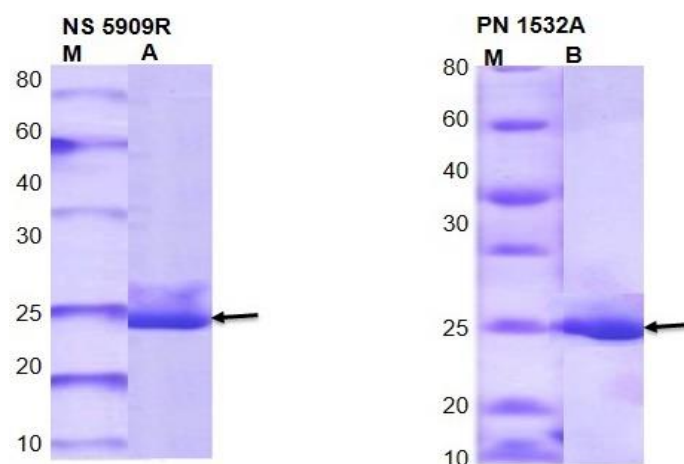


Figure 4.2: Purification of the recombinant GmAC protein fractions from the two *G. max* varieties under non-native denaturing conditions. The protein fractions were purified by SDS-PAGE on a 12% gel. M represents the standard molecular weight marker while A is variety NS 5909R known to be drought tolerant and B is variety PN 1532A known to be pest resistant. The arrows mark bands of the resolved GmAC protein fractions.

4.3.3 Characterization of the AC Activities of the Recombinant GmAC Protein Fractions

The cAMP-linked Enzyme Immunoassay System (Catalog # CA201; Sigma-Aldrich Inc., Missouri, USA) was used to evaluate the intrinsic AC activity of the purified and refolded recombinant GmAC protein fractions as well as further functionally characterizing such enzymatic activity *in vitro*. Following elution, the refolded, purified, recombinant GmAC from two types protein's AC activity was evaluated *in vitro* and in the presence of additional ionic and molecular elements known to naturally influence and/or affect AC activity. The cAMP levels produced by these different assessment conditions as shown in Figure 4.3 below.

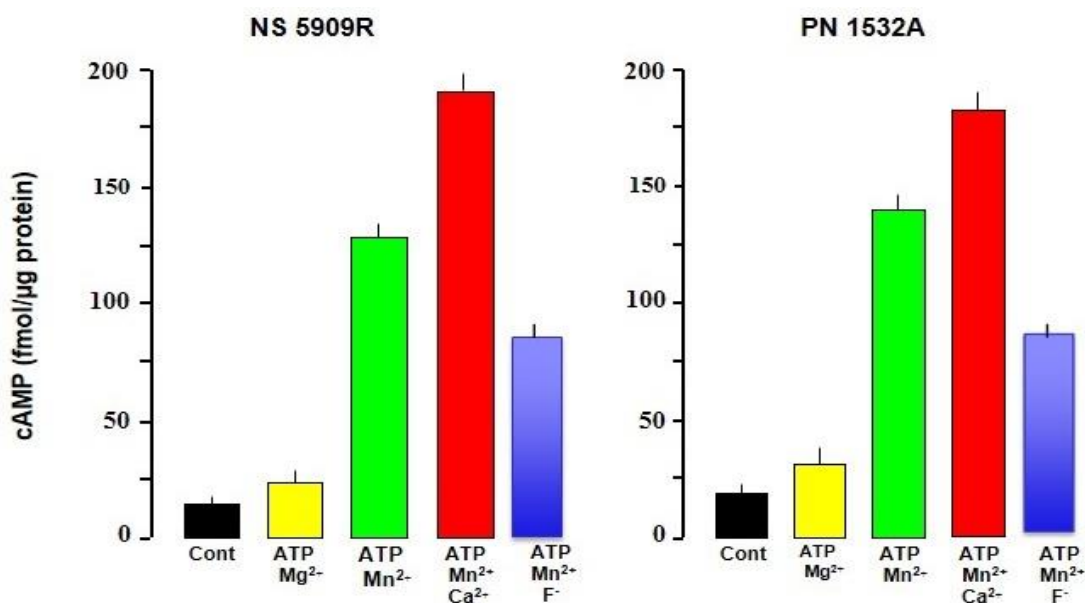


Figure 4.3: Characterization of the AC activity of the recombinant GmAC protein fractions isolated from the NS 5909R (1) or PN 1532A (2) *G. max* varieties. Reaction mixtures containing 10 μg of the refolded pure recombinant GmAC protein fraction, 50 mM Tris-Cl, pH: 8.0, 2 mM IBMX, 5 mM Mn²⁺, 1 mM ATP, and/or additional additives (calcium or fluoride ions) was prepared and incubated for 20 minutes at room temperature. cAMP levels generated from these reaction mixtures were quantified using a cAMP-specific enzyme immunoassaying system (Sigma Aldrich Inc., Missouri, USA All assays were carried out in triplicate sets (n =

3), and data then analyzed by one way analysis of variance (ANOVA). Error bars indicate standard errors of the mean.

4.4 Discussion

The reason behind the ability of *E. coli* BL21 (DE3) pLysS expression cells to produce inclusion bodies (IBs) of the studied recombinant GmAC protein fractions (Figure 4.1) was likely owed to the bacterial system's inability to support all of the post-translational modifications required by the GmAC protein fractions to properly fold and remain soluble and functional (Baneyx and Mujacic, 2004). This is because the recombinant GmAC protein fractions were produced in a bacterial (*E. coli*) micro-environment, which is substantially different from that of *G. max* (plant) (Yamaguchi and Miyazaki, 2014). Ideally, *E. coli* is prokaryotic while *G. max* is eukaryotic.

Notably, the presence of a C-terminal 6xHis tag on the pTrcHis2 cloning vector enabled successful affinity purification of the expressed GmAC protein fractions on the HisPur Ni-NTA affinity matrix because this negatively charged tag then ended up being part of the expressed recombinant GmAC proteins and subsequently affinity binding to the positively charged Ni-NTA matrix (Lindwall *et al.*, 2000; Stempfer *et al.*, 1996). Purification of the recombinant GmAC protein fractions was also easily made possible, thanks to the histidine tag's capability to function viably and effectively under denaturing and non-native experimental conditions (Holzinger *et al.*, 1996; Mihic and Harris, 1996). The purification method used somewhat caused the recombinant GmAC proteins to become denatured and rendering them physiologically inactive due to the presence of a high concentration of urea. As a result, it was critical to first recover physiologically active protein products through refolding (Yamaguchi and Miyazaki, 2014) and before proceeding with any further

experimental study. This refolding process, followed by elution, resulted in purified native protein fractions as is presented in Figure 4.2.

Functional analysis of the purified versions of the proteins *in vitro* showed that the GmAC fractions could generate cAMP from ATP (Figure 4.3). This capacity of the eluted recombinant GmAC protein fractions to produce cAMP from ATP was tested *in vitro* and also in the presence of other molecular and/or ionic components known to modulate ACs. Such components included the Mg^{2+} , Mn^{2+} , Ca^{2+} , and F^{-} ions. Previously, Mg^{2+} ion has been established as a required co-factor for most AC functions in several studies (Bradham, 1977), however, Steer and Levitzki, 1975 found out that the Mn^{2+} ion may actually better replace the Mg^{2+} ion as a co-factor and ultimately, activating the turkey erythrocyte soluble AC (sAC). In our case, the enzymatic activities of the recombinant GmAC protein fractions were evaluated using both the Mg^{2+} and Mn^{2+} metal ions as a co-factors, and it was discovered that the protein activities were indeed more dependent onto the Mn^{2+} metal ion than the Mg^{2+} metal ion (Figure 4.3). These outcomes were not surprising because Al-Younis *et al.* (2018) found that another Arabidopsis recombinant protein, AtKUP7 with AC activity, could generate 42.5 fmols of cAMP in the presence of Mn^{2+} and only 27.0 fmols of cAMP in the presence of Mg^{2+} . This outcome essentially signified GmAC as a typical soluble AC (sAC) because sACs utilize Mn^{2+} only as a co-factor while transmembrane ACs (tACs) utilize Mg^{2+} (Braun and Dods, 1975). In addition, the presence of Ca^{2+} and F^{-} ions caused an increase (~1.6 folds) and decrease (~1.4 folds) respectively of the activities of the GmAC protein fractions (Figure 4.3), further affirming GmAC as a sAC because sACs are specifically stimulated by Ca^{2+} and inhibited by F^{-} while the opposite is the case for tACs (Chen *et al.*, 2000).

Notably, no significant differences between the activities of the GmAC protein fractions from the NS 5909R or PN 1532A varieties were noted (Figure 4.3). Hence, not any probable variety-specific activity of the GmAC protein could, therefore, be ascertained and/or reported.

4.5 Conclusion

Based on findings of this chapter, the recombinant GmAC protein fractions from both of the two studied *G. max* varieties are sAC that utilize ATP as a substrate, Mn^{2+} as a co-factor, and whose catalytic activities are up-regulated by Ca^{2+} and down-regulated F^{-} .

4.6 Recommendation

In order to fully understand the function and/or regulatory mechanisms of this intriguing protein, further work on the GmAC either in intact cells or parts of living plant tissues needs to be undertaken. This could also perhaps reveal a possible variety-specific activity of the protein in *G. max*.

CHAPTER 5

COMPLEMENTATION TESTING OF THE RECOMBINANT GMAC PROTEIN FRACTIONS

Abstract

The levels of cyclic nucleotides in plants have been reported to be so much lower than those in animals, and this was the reason why the recognition of cyclic 3',5'-adenosine monophosphate (cAMP) as a messenger molecule in plant cells was delayed. Another and possibly more important reason for the reluctance in acceptance of cAMP as a messenger molecule in higher plants is that until quite recently, there was no genetic or molecular evidence of enzymes involved in its synthesis. As a result, *E. coli* was extensively investigated in order to better understand mechanisms behind its inducible lactose metabolic pathways. Plant scientists have already begun connecting this bacteria's system in the ability to ferment lactose by deleting the adenylyl cyclase (AC) catalytic domain from its only AC gene, resulting in *cyaA* mutants (SP850 hosts) that lack cAMP production, and as a result, unable to ferment lactose. Ideally, the system has now been taken advantage of and extensively used to execute a complementation test on putative AC candidate molecules from various organisms, whereby such candidates are used in the SP850 host to rescue lactose metabolism. Similarly, in order to check if the cloned GmAC protein from either of the two studied *G. max* varieties was an AC, we also tested it in this mutant host strain. As a result, the GmAC protein fractions from both the studied *G. max* varieties managed to rescue the SP850 mutant strain to ferment lactose, thereby demonstrating their enzymatic statuses as a functional ACs.

5.1 Introduction

The *cyaA* SP850 mutant provides a possible and very precise biochemical way of testing the AC activity of foreign recombinant proteins via a complementation phenomenon of the mutation. It has been demonstrated that when a *cyaA* mutant is transformed with a heterologous system, which encodes a putative AC gene, the AC gene will rescue the *cyaA* mutation (complementation) (Cotta *et al.*, 1998), allowing the mutant strain to now ferment lactose and thereby producing deep red or purplish colonies on MacConkey agar (Perlman and Pastan, 1969). Therefore, the specific goal for this chapter was to experimentally assess the probable AC catalytic activity of the annotated GmAC protein encoded by the Glyma.07G251000 gene through its possible complementation of the *cyaA* mutation in the *E. coli* SP850 mutant strain (Dole *et al.*, 2004). The mutant lacks the AC system required for cAMP generation and therefore, is unable to activate the catabolite activator protein, implying that its cAMP receptor protein-dependent promoter is inactive (Dole *et al.*, 2004). The mutant, especially while it is in its exponential phase, is a vigorous host for the recombinant expression systems of foreign proteins (Grossman *et al.*, 1998). It has also been found to be kanamycin-resistant (Cotta *et al.*, 1998) and pH-sensitive during growth (Ahmad and Newman, 1988).

A wide range of aerobic and facultative anaerobic bacteria have been demonstrated to require cyclic AMP for catabolite suppression (Pastan and Adhya, 1976; Botsford and Harman, 1992). cAMP is also involved in the responses and regulation of most signal transduction pathways that are important and essential for various organisms' adaptation and survival (Bahn *et al.*, 2007; Biswas *et al.*, 2007), as evidenced in Cyanobacteria, where its increased levels are a result of a response to environmental signals. cAMP has also

been practically demonstrated to transduce light signaling and cellular transduction systems in *Anabaena cylindrica* (Ohmori *et al.*, 1988).

In *E. coli*, cAMP is involved in the positive regulation of the *lac* operon, which is key to lactose fermentation (Ray *et al.*, 1987). Interestingly, despite the fact that most bacteria require cAMP for both signaling and survival, *E. coli* only has one AC, which is encoded by the *cyaA* gene (Côté *et al.*, 2010). Yang and Epstein discovered and purified this one-of-a-kind enzyme in 1983 that is crucial for lactose fermentation because it generates the most sought-after cAMP, which activates the *lac* operon (Cotta *et al.*, 1998). In 1992, Gross *et al* used site-directed (deletion) mutagenesis and genetic recombination to knock out this gene, resulting in the creation of the SP850 mutant strain (Shah and Peterkofsky, 1991). Due to the fact that this strain lacks AC activity, it is unable to metabolize carbon sources such as lactose, maltose, galactose, arabinose, and glycerol (Ullmann and Danchin, 1983; Moutinho *et al.*, 2001). When grown on MacConkey agar, this mutant and its wild-type have a very significant phenotypic difference, with the mutant appearing white or yellowish and the wild-type appearing deep red or purple (Cotta *et al.*, 1998). Nonetheless, the SP850 mutant strain has been observed to grow so slowly on glucose, fructose, and galactose sugars and is unable to ferment carbon sources like lactose, maltose, galactose, arabinose, and glycerol (Perlman and Pastan, 1969). The mutant is actually *lac* Z-deficient because it lacks the AC required for cAMP generation, and so is unable to activate the catabolite activator protein, implying that its cAMP receptor protein-dependent promoter is inactive (Dole *et al.*, 2004). If cAMP is available, it links to the bacteria's CRP to create the cAMP-CRP complex, which is bound to the CRP-binding area that is proximal to the *lac* promoter, thus boosting transcription in the *E. coli* cells (Kuo *et al.*, 2003). The *lac* promoter is essentially responsible

for controlling the expression of genes involved in lactose catabolism (Ullmann and Danchin, 1983).

5.2 Materials and Methods

Complementation tests were carried out in a mutant *E. coli* strain (the *cyaA* SP850) (Coli Genetic Stock Center, Yale University, Connecticut, USA) known to be defective in AC activity, to assess and determine the ability of the GmAC recombinant protein to generate cAMP from ATP and rescue the mutant.

5.2.1 Isolation and Purification of the pTrcHis2-TOPO:GmAC Expression Constructs from the *E. coli* EXPRESS BL21 (DE3) pLysS DUOs Cells

The pTrcHis2-TOPO:GmAC fusion expression construct for either of the two studied *G. max* varieties that was previously confirmed to be correctly and properly designed in Chapter 2 and maintained in the chemically competent *E. coli* EXPRESS BL21 (DE3) pLysS DUOs expression cells (Chapter 3) was extracted using the GeneJET Plasmid Miniprep Kit according to the manufacturer's instructions (Catalogue # K0503; Thermo Scientific Inc., New York, USA) and as previously outlined in section Chapter 3. The resultant pTrcHis2-TOPO:GmAC fusion expression construct was then used to transform chemically competent mutant SP850 host cells for undertaking the intended complementation testing.

5.2.2 Preparation of the Competent *E. coli* CyaA SP850 Cells

About 10 mL of the selective Luria-Bertani (LB) broth containing 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) sodium chloride, 0.5% (w/v) glucose, and 15 µg/ml kanamycin; pH: 7.0 were inoculated with a colony of the *E. coli* *cyaA* SP850 cells followed

by incubation overnight at 37°C at 200 rpm. The next morning, 20 ml of pre-warmed fresh LB broth supplemented with 15 µg/mL kanamycin were inoculated with 1 ml of the overnight culture and incubated at 37°C, shaking at 200 rpm up until an OD₆₀₀ of 0.5 was reached. Cells were then harvested by centrifugation at 4,000xg for 5 minutes at 4°C. The supernatant was discarded while the harvested cells were resuspended in 15 ml of an ice-cold Transformation Buffer 1 made-up of 30 mM potassium acetate, 50 mM manganese chloride, 100 mM rubidium chloride, 10 mM calcium chloride, and 15% (v/v) sterile glycerol; pH: 5.8 and incubated on ice for 90 minutes. The cells were harvested by centrifugation at 4,000xg for 5 minutes at 4°C. After centrifugation, the supernatant was discarded while the harvested cells were resuspended in 4 mL of ice-cold Transformation Buffer 2 containing 10 mM MOPS [3-(N-morpholino) propane sulfonic acid], 75 mM calcium chloride, 10 mM rubidium chloride, and 15% (v/v) glycerol; pH: 6.8. The resuspended cells were then aliquoted into 100 µL portions and kept on ice for further use in the subsequent steps.

5.2.3 Transformation of the Competent *E. coli* CyaA SP850 Cells

About 10 µL of the pTrcHis2-TOPO:GmAC fusion expression construct (Section 5.2.1) were added into an ice-cold 1.5 ml microfuge tube containing 100 µl of the chemically competent SP850 *E. coli* mutant cells (Section 5.2.2). The mixture was gently mixed through swirling and then incubated on ice for 20 minutes. The mixture was heat shocked through incubation at 42°C for 90 seconds on a dry bath heating block and immediately brought back for incubation on ice for 5 minutes. The transformation mixture was supplemented with 500 µL of the super optimal broth (SOC) medium and incubated at 37°C for 90 minutes in an orbital shaker at 200 rpm. This step was undertaken so as to allow cells to express the β-lactamase enzyme, which during and later on in the subsequent selection process would detoxifies ampicillin and kanamycin. After incubation, 100 µl aliquots of the transformation mixture

were then used to spread-plate LB agar plates supplemented with 15 µg/ml kanamycin and 100 µg/ml ampicillin. The inoculated plates were incubated at 37°C overnight for the generated cell colonies to be used in the next step.

5.2.4 Cell Culturing and Phenotypic Analysis on MacConkey Agar

About 10 ml of the LB broth were inoculated with a colony of a transformed *E. coli cyaA* SP850 cell from section 5.2.3 above, followed by incubation at 37°C and shaking at 200 rpm. At an OD₆₀₀ of 0.5, cells were split into two equal portions of 5 mL each. One portion was then supplemented with 0.5 mM IPTG (to induce protein expression) while the other portion was left un-induced (control). Alongside this culturing process, a MacConkey agar plate containing 2% (w/v) peptone, 1% (w/v) lactose, 0.5% (w/v) bile salts, 0.0075% (w/v) neutral red, and 1.5% (w/v) agar was prepared and divided into four segments using a permanent marker, labeled A to D. Segment A was streaked with the mutant SP850 cells, segment B with the un-induced transformed SP850 cells, segment C with the induced transformed SP850 cells, and segment D with the wild-type *E coli* cells. The streaked plate was inverted and incubated at 37°C for 40 hours before its segments were then visually analysed for different phenotypic characteristics (MacConkey, 1905). A deep red or purple colour on the transformed SP850 mutant cells would mean a positive AC activity for the cloned and recombinantly expressed GmAC protein (Cotta *et al.*, 1998).

5.3 Results

To confirm the cyclase activities of the recombinant GmAC protein fragments isolated from the two *G. max* varieties, the recombinant protein fragments were expressed in a *cyaA* mutant SP850, followed by assessment of their ability to rescue this mutant host. Here, the mutant cells were transformed with the pTrcHis2TOPO fusion expression vector carrying the Gm-AC gene fragment (pTrcHis2- TOPO:Gm-AC) from two varieties (Gm AC1 and Gm AC2). The Gm-ability AC's to produce cAMP on a differential and selective MacConkey agar media was then examined, as shown in Figure 5.1 below.

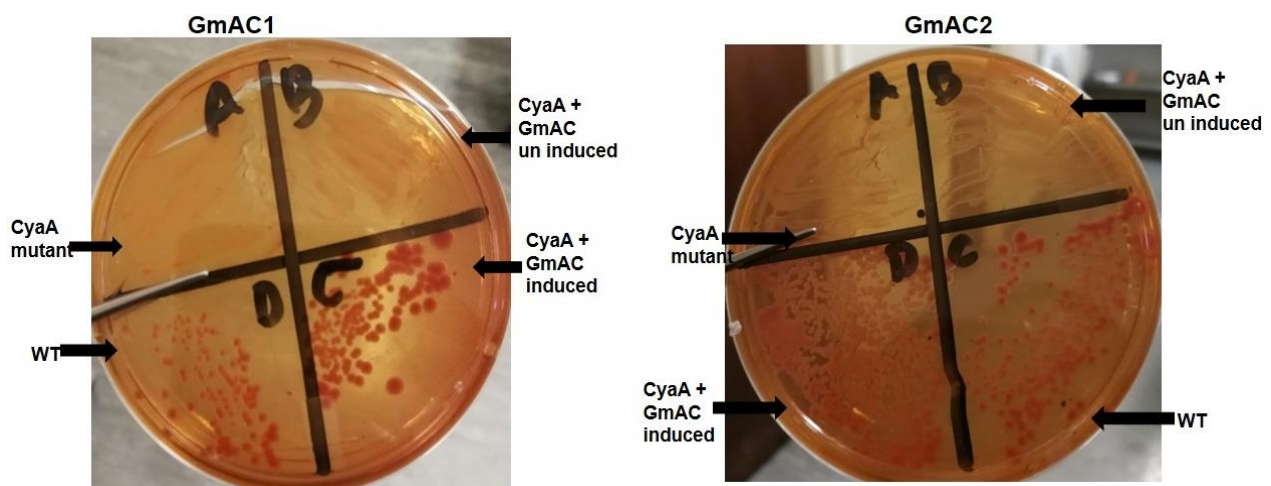


Figure 5.1: Testing for the ability of the GmAC protein fractions from the two *G. max* varieties to rescue *cyaA* mutation in SP850 cells. For both protein fragments from the two studied *G. max* varieties, (A) mutant SP850 cells and (B) un-induced transformed SP850 cells produced colourless colonies while the (C) induced transformed SP850 cells and (D) wild-type cells yielded reddish colonies.

5.4 Discussion

The *cyaA* SP850 mutant strain was produced by using a P1 background transduction technique to transfer a whole gene deletion of the *cya* gene from the SP793 strain to the CA8000 strain (Shah and Peterkofsky, 1991). It was then demonstrated in this experiment that the freshly created SP850 strain exhibited all of the features associated with the lack of AC activity. The strain was later recommended for use as a clean *cya* deletion mutant before its deposition into the E. Yale University's *E. coli* Genetic Stock Center. Through a complementation phenomenon of the mutation, the *cyaA* SP850 mutant strain has supplied a highly valuable system to test for the possible AC activities of foreign recombinant proteins in today's modern science system. In genetic terms, the complementation test is used to assess whether two recessive mutations that cause the same phenotype are located on the same gene (Hanoune and Defer, 2001).

In order to check and establish if the cloned and expressed GmAC protein fragments from the two studied *G. max* varieties had any AC activity, the recombinant protein fragments were expressed in a *cyaA* mutant, SP850. As a result, both protein fragments rescued the *cyaA* mutation of the mutant SP850 host to now metabolize lactose sugar just like the wild-type *E. coli* does (Figure 5.1). This, therefore, unequivocally demonstrated that the expressed protein fragments from both the *G. max* varieties were indeed functional ACs with the capability of generating cAMP from ATP. However, no noticeable differences between the protein fragment from the NS 5909R variety and that of the PN 1532A variety were noted (Figure 5.1) to proclaim any probable variety-specific activity of the GmAC protein.

This result or outcome was typically consistent with ten of the eleven currently known plant ACs as is summarised here. It has been reported , that a PSiP protein involved in the polarized development of pollen tubes in *Zea mays* can rescue the same SP850 *cyaA* mutant host (Moutinho *et al.*, 2001), while reported that a phenotypic manifestation of a magenta colour of the recombinant cells on MacConkey lactose agar after cloning the AC fragment domain of a pentatricopeptide (PPR) protein from *Arabidopsis thaliana* into the SP850 mutant (Ruzvidzo *et al.*, 2013). Another plant AC, HpAC1 from *Hippeastrum hybridum*, was found to be capable of rescuing the AC deficiency in SP850 cells through complementation (Świeżawska *et al.*, 2014) while (Al-Younis *et al.*, 2015; Al-Younis *et al.*, 2018) also demonstrated the rescue of the *E. coli* SP850 strain by two *Arabidopsis* recombinant proteins, the AtKUP7 and AtKUP5. The similar technique was done to show that the SP850 cells could be successfully complemented by a MpCAPE recombinant protein from a lower plant, *Marchantia polymorpha* was (Kasahara *et al.*, 2016). Similarly, the *Arabidopsis* recombinant fragments, AtCIAP and AtLRRAC1 could rescue *E. coli* AC deficiency in the SP850 mutant strain (Chatukuta *et al.*, 2018; Bianchet *et al.*, 2019). Finally, the ZmRPP13-LK3 protein fragment from maize used to rescue the *cyaA* mutation in the SP850 strain (Yang *et al.*, 2021).

In this chapter, we assess the expression of AC-containing gene fragments of the Glyma.07G251000 gene isolated from two *G. max* varieties in SP850 cells and show their ability to rescue the *cyaA* mutation in this mutant host

5.5 Conclusion

The findings in this chapter demonstrated that the recombinant GmAC protein fragment from either of the two studied varieties of *G. max* is a true functioning AC, capable of converting

ATP to cAMP. The protein is the first ever of its kind to be discovered in soybean specifically and legumes in general.

5.6 Recommendation

Despite the fact that the GmAC protein from either of the two studied varieties has been established as a *bone fide* higher plant AC molecule in this chapter, its exact or actual physiological roles in plants, particularly the very sought variety-specific ones, have yet to be fully explored. Therefore, further work on this novel protein molecule was herein strongly recommended.

CHAPTER 6

EXPLORING THE INVOLVEMENT OF THE GMAC PROTEIN FRACTIONS IN STRESS RESPONSE

Abstract

The cyclic nucleotide molecule, cAMP (3',5'-cyclic adenosine monophosphate), is now recognized as a key signaling molecule in plants, participating in a variety of molecular processes such as sensing and responding to biotic and abiotic environmental challenges. Plant AC discoveries and cAMP production strengthened the need to undertake further and advanced research on cAMP-dependent signaling systems in plants. After successfully establishing GmAC, isolated from either of the two studied *G. max* varieties, as a *bona fide* AC molecule in Chapters 3 - 5, we then further sought to characterize this protein molecule with respect to its possible role in stress response. Specifically, the protein was tested for its possible capability to confer stress tolerance to *E. coli* cells expressing it under salt and oxidative stress. To achieve this, EXPRESS BL21 (DE3) pLysS DUOs cells harboring the pTrcHis-TOPO:GmAC fusion construct from either of the two studied *G. max* cultivars were cultured in LB media supplemented with 600 mM NaCl or 0.6 mM H₂O₂, followed by evaluation of cell growth by liquid and spot culture assays. For cells growing in liquid media, growth was measured through culture absorbances at OD₆₀₀ nm at 2-hour intervals for 24 hours while for cells growing on solid media, growth was through colony diameter measurements after 24 hours of incubation. Findings showed that the GmAC protein could indeed confer resistance to both 600 mM NaCl and 0.6 mM H₂O₂, thus suggesting a key role for it in resistance to both biotic and abiotic environmental stress factors.

6.1 Introduction

To adjust to harsh environmental conditions, plants synthesize various kinds of signal molecules to regulate gene expression and physiological processes. Cyclic nucleotide monophosphates (cNMPs) are one such group known to be well-conserved second messengers in plants, involved in many biological processes (Yang *et al.*, 2021, Moutinho *et al.*, 2001; Ruzvidzo *et al.*, 2013; Ito *et al.*, 2014; Swieżawska *et al.*, 2014; Al-Younis *et al.*, 2015; Marondedze *et al.*, 2015; Chatukuta *et al.*, 2018). Cyclic adenosine monophosphate (cAMP) is one type of cNMPs that plays important roles in such many processes, including lipid metabolism, stomatal opening, cell division and differentiation, and stress responses (Ehsan *et al.*, 1998; Gottig *et al.*, 2009; Thomas *et al.*, 2013; Alqurashi *et al.*, 2016).

In *Arabidopsis thaliana*, salicylic acid (SA) and jasmonic acid (JA) can induce cAMP synthesis, which then functions as a ligand to activate cyclic nucleotide gated channel 2 (CNGC2), resulting in an influx of extracellular Ca^{2+} . CNGCs, which are cAMP- and voltage-gated channels, have been shown to play a vital role in regulating plant development and stress responses (Yang *et al.*, 2021). Heat can increase the concentration of cytosolic cAMP in *Arabidopsis*, which then leads to an influx of Ca^{2+} ions, mediated by CNGC and in turn, enhancing the expression of heat-shock proteins (HSPs) that facilitate tolerance to heat stress (Liang *et al.*, 2022). cAMP is mainly synthesized by adenylyl cyclases (ACs), which catalyze the conversion of ATP to cAMP and pyrophosphate.

Plant response to salt stress revealed a link between cAMP, ROS, and ion homeostasis (Serrano *et al.*, 1999). ROS production in salt-stressed roots has been linked to mitochondrial dysfunction as well as activation of NADPH oxidases, according to several

studies (Sharma *et al.*, 2012). Furthermore, when a cell is exposed to salt, Na⁺ inflow produces a considerable loss of cytosolic K⁺, which can lead to major metabolic changes (Blanco *et al.*, 2020). Plants are also able to withstand biotic stressors through a variety of mechanisms, including pathogen-associated molecular patterns (PAMPs), which activate immune and plant resistance proteins (Moustafa-Farag *et al.*, 2019). Pathogens can produce virulence factors encoded by avirulence (avr) genes in plant cells, which decrease pattern-triggered immunity (PTI). The second layer of defense, effector triggered immunity, is activated when an effector is recognized by intracellular receptors encoded by resistance genes (ETI). ETI defense responses are often greater than PTI defense responses, and they frequently end in the hypersensitive response (HR), a type of programmed cell death (PCD) that occurs at the infection site with the goal of limiting pathogen infection (Künstler *et al.*, 2016). Antimicrobial phytoalexins, as well as SA, JA and pathogenesis-related (PR) proteins, are upregulated both locally and systemically in diseased tissues. Although there is no literature on the relevance of mitochondrially-generated cAMP in plants or soybeans in particular, earlier research had focused onto the effects of similar signaling systems in animals (Valsecchi *et al.*, 2014). Glycolysis and mitochondrial respiration were reported to be regulated by SACs in mammals (Acin-Perez *et al.*, 2009), and ACs were also reported to be involved in the regulation of pathogens such as *Candida albicans* and *Pseudomonas aeruginosa* (Hall *et al.*, 2010; Topal *et al.*, 2012).

Previous studies suggest that specific cyclic nucleotide signatures generated in response to both the biotic and abiotic stress factors act as messengers in signaling cascades that critically depend on CNGCs (Nawaz *et al.*, 2014). cAMP and ACs have been shown to play key roles in biotic (Alqurashi *et al.*, 2016) and abiotic (Gehring, 2010) stresses. Therefore, in an attempt to determine if the GmAC protein, isolated from either of the two studied G.

max varieties, had some functional involvement in any of the 2 stress factors, we cloned the AC-containing XP_003529590 gene fragments (GmAC fractions) into the EXPRESS BL21 (DE3) pLysS DUOs cells, followed by their growth and assessment of their growth responses under salt (abiotic) and peroxide (biotic) conditions.

6.2 Material and Methods

6.2.1 Testing for the Ability of GmAC Protein Fractions to Confer Stress Tolerance to *E. coli*

Liquid and spot culture assays were carried out to test and determine the possible ability of GmAC to confer stress tolerance to *E. coli* cells.

6.2.1.1 Liquid Culture Assaying

A total of 20 ml of 2YT broth media containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol and supplemented with 600 mM NaCl (for salt stress) or 0,6 mM H₂O₂ (for oxidative stress), were inoculated with 1 ml of the EXPRESS BL21 (DE3) pLysS DUOs cells harboring the pTrcHis-TOPO:GmAC fusion construct of either of the two studied *G. max* varieties (Chapter 3). Control cultures were not supplemented with anything. The prepared cultures were incubated at 37°C in an orbital shaker at 200 rpm. At an OD₆₀₀ of 0.5, the cell culture was split into two equal portions of 20 ml each. In the first cell culture, 1 mM IPTG was added (to induce protein expression) while the second culture was left untreated (control). The two cultures were then left to grow for a further 24 hours at 37°C at 200 rpm while at every 2-hour point, 1 ml of the culture was collected and its OD₆₀₀ measured using

a Helios Epsilon spectrophotometer (Thermo Scientific Inc., Madison, USA). The recorded readings were then plotted into graphs and analysed comparatively.

6.2.1.2 Spot Culture Assaying

The same cells in Section 6.2.1.1 above (EXPRESS BL21 (DE3) pLysS DUOs cells transformed with either of the pTrcHis-TOPO:GmAC fusion constructs), at an OD₆₀₀ of 0.5, were similarly split in two portions of the induced and non-induced cultures. Each of the split cultures was then serially diluted to 10⁰, 10⁻¹, 10⁻² and 10⁻³ folds using pure 2YT media, followed by spotting of 10 µl of each diluent onto nutrient agar plates containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol, and supplemented with either 600 mM NaCl (salt stress) or 0.6 mM H₂O₂ (oxidative stress) or no supplementation at all (control). The prepared plates were then incubated in an Incubat bench-top incubator (JP Selecta Inc., Barcelona, Spain) at 37°C for 24 hours. Thereafter, the diameter of each cell colony was measured in mm with a ruler, to ascertain the extent of cell growth, and then a comparative analysis of the cell colony sizes undertaken.

6.3 Results

From the undertaken liquid culture assay and for both the studied *G. max* varieties, in ordinary liquid media not exposed to any stress condition, cells expressing the GmAC protein and those not expressing it grew relatively the same, reaching a maximum growth density of around 1.25 nm (OD₆₀₀) after about 18 hours of incubation (Figure 6.1A1). However, the pattern of growth was noticeably different for cells growing in media with stress where, in media with salt stress, cells expressing the GmAC protein reached maximum growth after 20 hours of incubation with a culture density of around 1.0 nm compared to

cells not expressing the protein, which had a maximum growth density of 0.6 nm after a whole 24-hour period of incubation (Figure 6.1AII). In media with oxidative stress, cells expressing the GmAC protein also reached maximum growth after 20 hours of incubation with a culture density of around 1.06 nm compared to cells not expressing the protein, which literally struggled to grow for the entire 24 hours of incubation, with a mere maximum growth of around 1.99 nm (Figure 6.1AIII).

With regards to spot culture assaying and for both the studied *G. max* varieties, agar plates with no stress condition, similarly, showed no noticeable differences in colony sizes (~7.0 mm) between the *E. coli* cells expressing and not expressing the GmAC protein and such an observation was also made for all cellular dilution factors used in the experiment and after the complete incubation period of 24 hours (Figure 6.1BI). However, for agar plates with stress conditions, the picture was different, whereby all colony sizes, at various dilution factors, of cells expressing the GmAC protein, were noticeably larger (~5.0 mm) than those of cells not expressing the protein (~2.0 mm) whether growing under salt or oxidative stress (Figure 6.1BII-III).

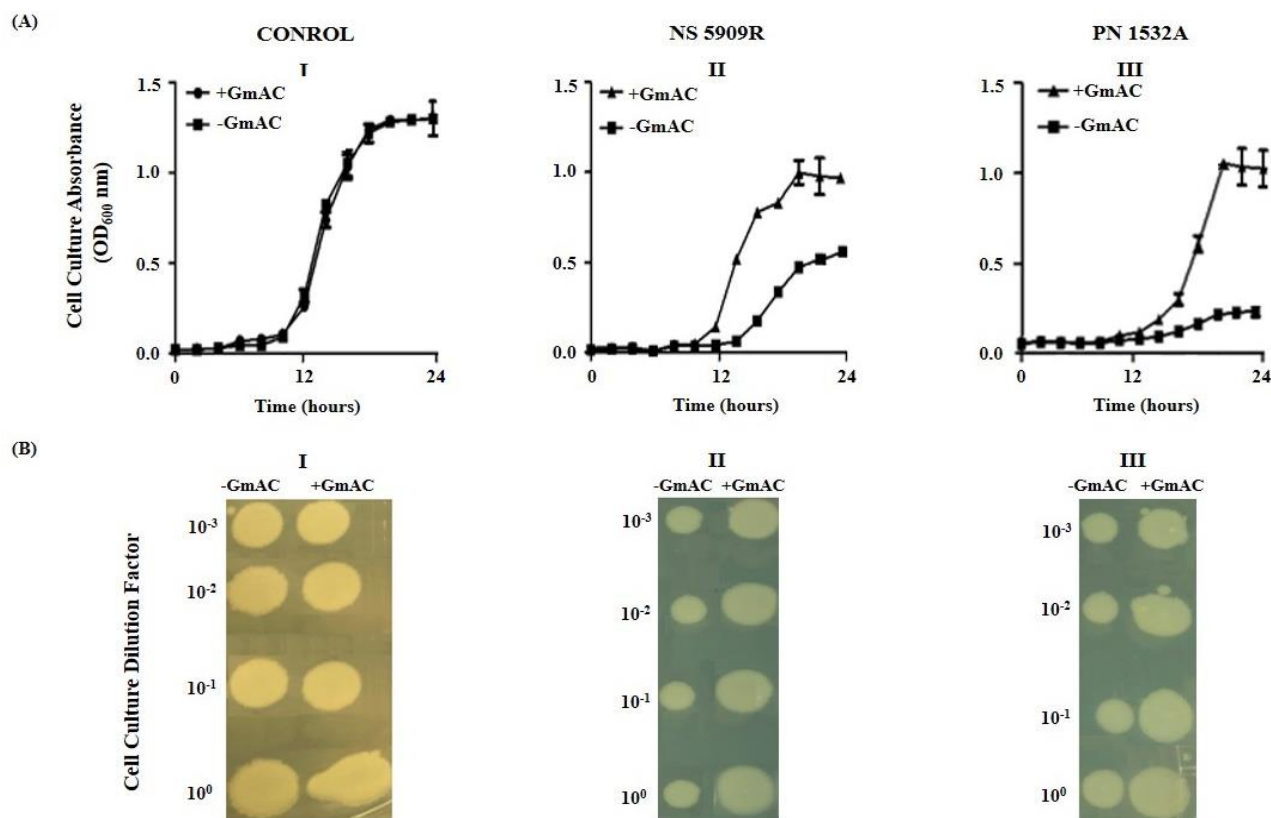


Figure 6.1: Determining the role of GmAC in stress response. *E. coli* cells expressing (+GmAC) and not expressing (-GmAC) the GmAC protein were cultured in (A) liquid and (B) solid media that was either ordinary (I) or supplemented with 600 mM NaCl for salt stress and (II), 0.6 mM H₂O₂ for oxidative stress (III), followed by evaluation of their growth patterns. For cells growing in liquid media, growth was measured through culture absorbances at OD₆₀₀ nm at 2-hour intervals for 24 hours while for cells growing on solid media, it was through colony diameter measurements after 24 hours of incubation.

6.4 Discussion

cAMP has been regarded as a significant signaling molecule in many biological processes, including growth and differentiation, photosynthesis, and stress defense, as more ACs have been identified in plants and more accurate techniques for quantification have become available (Yang *et al.*, 2021). The activation of numerous processes at the gene, transcript, and protein levels, coupled by the activity of signaling messengers, is required to develop plant responses to environmental stimuli. The cellular metabolism of stressed plants

undergoes a remarkable re-organization, allowing stress adaptation. The development of oxidative stress and the stimulation of stress responsive signaling pathways are early alarm stages of plant abiotic stress response. Cells encounter new recovering homeostasis after the acclimation period, with the manufacture of stress protective chemicals, at the expense of cellular energy (Blanco *et al.*, 2020). In this setting, cAMP may operate as a stress sensor and/or modulator of cellular metabolism, mostly but not exclusively through its action on ion channels and the regulation of ion fluxes that results.

Plant salinity tolerance has been noted to be improved by cAMP in *Arabidopsis*, which causes the deactivation of voltage-independent nonselective channels, decreasing Na⁺ influx (Al-Younis *et al.*, 2015). Tolerance to aluminum in wheat necessitates a cAMP-dependent outward rectifying K⁺ current, which allows malate outflow, and chelating the harmful element. The recent identification and description of two K⁺ uptake permeases, AtKUP5 and AtKUP7, in *Arabidopsis thaliana* further highlighted the crucial relationship between salt stress and cAMP synthesis. Both permeases have a dual function, with an AC catalytic domain that is active and another one, which regulates ion fluxes (Blanco *et al.*, 2020). The identification of ACs engaged in plant pathogen response, further validated the role of cAMP in plant immunity. For instance, the bacterial toxin *tabtoxinine lactam*, produced by *Pseudomonas syringae* pv. *Tabaci* and causing necrotic lesions in *Nicotiana benthamiana*, was found to be silenced when an NbAC gene, producing an AC, was expressed (Ito *et al.*, 2014). Many studies have also further revealed the importance of endogenous cAMP in plant defense responses, resulting in increased transcript levels of defense-related genes. Several investigations have found that infection sites produce increased cAMP concentrations. These were seen in *Phaseolus vulgaris* (Bolwell, 1992) and *Medicago sativa* (Cooke *et al.*, 1994) treated with *Verticillium alboatrum* glycoprotein,

as well as *Arabidopsis* exposed to *Verticillium dahlia* toxins (Jiang *et al.*, 2005). The increased cAMP concentrations in all of these cases, have resulted in better disease resistance in the plants. The importance of cAMP in increasing plant defense against pathogen attack cannot be overstated. Wounding stress was also demonstrated to result in a nearly five-fold rise in the concentration of cAMP in *A. thaliana* leaves (Van Damme *et al.*, 2014).

Therefore, in order to further augment the involvement of cAMP in stress response, we used liquid and spot culture assays to demonstrate the role of Glyma.07G251000 and its gene product GmAC in this key process. As is shown in Figure 6.1, GmAC protein fragment from either of the studied *G. max* varieties conferred resistance to *E. coli* cells expressing it under salt stress, which mimicked drought and oxidative stress that mimics pathogenesis. Such an AC function was closely related to functions of some of the previously identifies proteins in the higher plant family. For instance, a *Nicotiana benthamiana* protein, NbAC, was found to play a key role in *tabtoxinine*- β -*lactam*-induced cell deaths during wildfire diseases (Ito *et al.*, 2014) while a *Hippeastrum hybridum* protein, HpAC1 was also found to be involved in stress signaling (Świeżawska *et al.*, 2014). Furthermore, two *A. thaliana* K⁺-uptake permeases, AtKUP5 and AtKUP7, were both found to be responsible for K⁺ fluxes across cells during high salt challenges (Al Younis *et al.*, 2015; Al-Younis *et al.*, 2018) and lastly, a clathrin assembly protein, AtCAP, and a leucine-rich repeat protein, AtLRR, from *A. thaliana*, were both found to affect responses to pathogens (Chatukuta *et al.*, 2018; Bianchet *et al.*, 2019).

Therefore, the research findings presented herein, further demonstrate and endorse the fact that cAMP has a key function in the control and/or regulation of both biotic and abiotic stress

factors in *G.max*. However, no significant differences between the activities of the two studied GmAC protein fractions were observed. Therefore, no probable variety-specific activity of the GmAC protein could be stated.

6.5 Conclusion

In conclusion, our research demonstrates that the GmAC protein plays a critical role in conferring stress tolerance. This is particularly true for biotic and abiotic stress factors.

6.6 Recommendation

Despite the fact that GmAC has now been clearly implicated in stress response, the molecular mechanisms by which this protein is involved are still not known. Therefore, it is hereby recommended that more and further research be undertaken to establish exactly how this protein is activated and gets involved in these key cellular processes of plants.

CHAPTER 7

OVERALL DISCUSSION AND RECOMMENDATIONS

7.1 Overall Discussion

In this research study, fractions of the GmAC protein annotated to be an adenylyl cyclase (AC), were prepared from two varieties of *Glycine max* (NS 5909R known to be drought tolerant while PN 1532A known to be pest resistant), followed by assessment of their possible variety-specific activities. The prepared protein fractions were comparatively assessed for endogenous activity in Chapter 3, *in vitro* activity in Chapter 4, complementation activity in Chapter 5 and functional role in stress response in Chapter 6. Across all these assessments and evaluations, it was found out that there were no significant differences in the functional activities of the two studied GmAC protein fractions and thus no particular variety-specific activity too between the two studied protein fractions.

After realizing this overall outcome, we then decided to take a closer look at the AC catalytic center (Figure 7.1A) of these two protein fractions, that is primarily responsible for the generation of cAMP from ATP. The stated closer look involved chemically sequencing the AC catalytic center (Inqaba Biotechnological Sciences (Pretoria, RSA) followed by a concerted analysis of the resultant sequences. As is shown in Figure 7.1B below, the only difference detected between the two studied protein fractions was on amino acid number 3 in the catalytic center, whereby the NS 5909R protein fraction had an E (glutamic acid) while the PN 1532A fraction had a D (aspartic acid). Notably, this observed difference was so much thrilling in that it did not appear to be unusual because of two specific reasons. Firstly, in the AC catalytic center, at position 3 (Figure 7.1A), either an E or D residue is expected,

which are both acidic amino acids, to perform similar or related functions (Gehring, 2010). Secondly, these two residues are also commonly found in other previously identified and experimentally confirmed plant ACs (Figure 7.1 C & D), whereby E is found in AtCIAP, AtLRRAC1 and NbAC (Chatukuta *et al.*, 2018; Bianchet *et al.*, 2019; Ruzvidzo *et al.*, 2019; Ito *et al.*, 2014;) while D is found in AtPPRAC, AtKUP7, AtKUP5, AtNCED3, ZmPSiP and ZmRPP13-LK3 (Ruzvidzo *et al.*, 2013; Al-Younis *et al.*, 2015, 2018, 2021; Moutinho *et al.*, 2001; Yang *et al.*, 2021).

(A) AC catalytic motif

- [RKS]X[DE]X{9,11}[KR]X{1,3}[DE]-
 1 3 14

(B) GmAC amino acid sequence

--LEVIKHQ**K****E****D**DRKALAAS**L****K****S**DILVMVSELLSMGLEANVSLEDLDRKEDIPQ-- NS 5909R
 --LEVIKHQ**K****E****D**DRKALAAS**L****K****S**DILVMVSELLSMGLEANVSLEDLDRKEDIPQ-- PN 1532A

(C)

NS 5909R and other experimentally confirmed plant ACs:

NS5909R - **K****E****D**DRKALAAS**L**-**K****S**-**D** -
 AtCIAP - **K****W****E**IFEDDYRCFDR**K**-**D** -
 AtLRRAC1 - **K****M****E**KVVRLLEHHV**K****H****I****E** -
 NbAC - **R****L****E**VIKRQKDEKR**K**--**E** -

(D)

PN 1532A and other experimentally confirmed plant ACs:

PN1532A - **K****E****D**DRKALAAS**L**--**K****S**--**D** -
 AtPPRAC - **K****F****D**VVISLGEK**M****Q**-**R****L**--**E** -
 AtKUP7 - **S****F****D**VEALEVPGAP-**R****N**--**D** -
 AtKUP5 - **S****F****D**VDALEIPGT**Q**-**K****N**--**E** -
 AtNCED3 - **S****Y****D**VVSKPYLK**Y****F**-**R****F****S****P****D** -
 ZmPSiP - **S****V****D**VFAIVGVGG--**K****T**--**D** -
 ZmRPP13-LK3 - **K****V****D**NARKMMTEKE**E****K****I****W****E** -

Figure 7.1: Sequence analysis of the AC catalytic center. **(A)** The 14 amino acid AC search motif derived from the catalytic centres of annotated and experimentally tested nucleotide cyclases (NCs). The residue forming hydrogen bonding with a purine at position, 1 is highlighted in red; the residue conferring substrate specificity in position 3, is highlighted in blue; while the amino acid in position 14, stabilizing the transition state from ATP to cAMP, is highlighted in red. The amino acid [DE] at 1–3 residues downstream from position 14, participates in Mg²⁺/Mn²⁺-binding and is coloured green. **(B)** Alignment of the GmAC protein fragments from the NS 5909R and PN 1532A varieties of *G. max*, showing the only single amino acid residue difference in

their catalytic centers at position 3 (marked with the two red arrows). **(C)** Alignment of the GmAC protein fragment from the NS 5909R variety of *G. max*, showing its similarity at position 3 of its catalytic center with the other previously confirmed plant ACs (marked with the two red arrows). **(D)** Alignment of the GmAC protein fragment from the PN 1532A variety of *G. max*, showing its similarity at position 3 of its catalytic center with the other previously confirmed plant ACs (marked with the two red arrows).

From this study, it can be confidently concluded that ACs are highly conserved molecules in *G. max*, whose role is to generate cAMP that is critically required by other cellular processes essential to plants. This also suggests that other genes, and not ACs, are the ones responsible for determining varieties and/or cultivars in *G. max* although the functions of such genes in their specific roles will be strictly regulated and/or controlled by cAMP. It is also worth noting that considering the key and essential role of cAMP in most cellular, physiological and biochemical processes of plants, it is very reasonable that ACs should be highly conserved and remain so while other genes are variable.

7.2 Recommendations

It is strongly advised that future studies of the GmAC protein include testing its functions in living plant tissues, followed by thorough bioinformatics correlational expression analysis in conjunction with the other known genes involved in plant signaling. In order, to learn more about the gene at a total plant level, studies employing GmAC knockouts or knockdowns must be conducted. Furthermore, exceptional and functionally uncharacterized putative protein candidates must be functionally described in a manner similar to how the GmAC was characterized in this investigation. Therefore, gaining a better understanding of their precise physiological, biological, and expressional roles in plants, especially in crucial cellular processes like growth, development, and responses to stressful environmental factors. This

could serve as a basis for developing strategies and countermeasures that could assist crops withstand the current threat of environmental extremes brought on by climate change.

REFERENCES

- Acin-Perez, R., Salazar, E., Kamenetsky, M., Buck, J., Levin, L.R. and Manfredi, G., 2009.** Cyclic AMP produced inside mitochondria regulates oxidative phosphorylation. *Cell Metabolism*, 9(3), pp.265-276.
- Agarwal, P.K., Agarwal, P., Reddy, M.K. and Sopory, S.K., 2006.** Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. *Plant Cell Reports*, 25(12), pp.1263-1274.
- Aghaei, K., Ehsanpour, A.A., Shah, A.H. and Komatsu, S., 2009.** Proteome analysis of soybean hypocotyl and root under salt stress. *Amino Acids*, 36(1), pp.91-98.
- Ahanger, M.A., Tittal, M., Mir, R.A. and Agarwal, R.M., 2017.** Alleviation of water and osmotic stress-induced changes in nitrogen metabolizing enzymes in *Triticum aestivum* L. cultivars by potassium. *Protoplasma*, 254(5), pp.1953-1963.
- Ahmad, D. and Newman, E.B., 1988.** A deficiency in cyclic AMP results in pH-sensitive growth of *Escherichia coli* K-12. *Journal of Bacteriology*, 170(8), pp.3443-3447.
- Asgher, M., Qamar, S.A., Bilal, M. and Iqbal, H.M., 2020.** Bio-based active food packaging materials: Sustainable alternative to conventional petrochemical-based packaging materials. *Food Research International*, 137, p.109625.
- Alqurashi, M., Gehring, C. and Marondedze, C., 2016.** Changes in the *Arabidopsis thaliana* proteome implicate cAMP in biotic and abiotic stress responses and changes in energy metabolism. *International Journal of Molecular Sciences*, 17(6), p.852.
- Al-Younis, I., Wong, A. and Gehring, C., 2015.** The *Arabidopsis thaliana* K⁺-uptake permease 7 (AtKUP7) contains a functional cytosolic adenylate cyclase catalytic centre. *FEBS Letters*, 589(24), pp.3848-3852.
- Al-Younis, I., Wong, A., Lemtiri-Chlieh, F., Schmöckel, S., Tester, M., Gehring, C. and Donaldson, L., 2018.** The *Arabidopsis thaliana* K⁺-uptake permease 5 (AtKUP5) contains a

functional cytosolic adenylate cyclase essential for K⁺ transport. *Frontiers in Plant Science*, 9, p.1645.

Bahn, Y.S., Molenda, M., Staab, J.F., Lyman, C.A., Gordon, L.J. and Sundstrom, P., 2007. Genome-wide transcriptional profiling of the cyclic AMP-dependent signaling pathway during morphogenic transitions of *Candida albicans*. *Eukaryotic Cell*, 6(12), pp.2376-2390.

Baneyx, F. and Mujacic, M., 2004. Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology*, 22(11), pp.1399-1408.

Bhatwa, A., Wang, W., Hassan, Y.I., Abraham, N., Li, X.Z. and Zhou, T., 2021. Challenges associated with the formation of recombinant protein inclusion bodies in *Escherichia coli* and strategies to address them for industrial applications. *Frontiers in Bioengineering and Biotechnology*, 9, p.65.

Basal, O. and Szabó, A., 2020. Physiomorphology of Soybean as Affected by Drought Stress and Nitrogen Application. *Scientifica*, 2020, p.6093836.

Bianchet, C., Wong, A., Quaglia, M., Alqurashi, M., Gehring, C., Ntoukakis, V. and Pasqualini, S., 2019. An *Arabidopsis thaliana* leucine-rich repeat protein harbors an adenylyl cyclase catalytic center and affects responses to pathogens. *Journal of Plant Physiology*, 232, pp.12-22.

Biswas, S., Van Dijck, P. and Datta, A., 2007. Environmental sensing and signal transduction pathways regulating morphopathogenic determinants of *Candida albicans*. *Microbiology and Molecular Biology Reviews*, 71(2), pp.348-376.

Blanco, E., Fortunato, S., Viggiano, L. and de Pinto, M.C., 2020. Cyclic AMP: A polyhedral signalling molecule in plants. *International Journal of Molecular Sciences*, 21(14), p.4862.

Bolwell, G.P., 1992. A role for phosphorylation in the down-regulation of phenylalanine ammonia-lyase in suspension-cultured cells of French bean. *Phytochemistry*, 31(12), pp.4081-4086.

- Botsford, J.L. and Harman, J.G., 1992.** Cyclic AMP in prokaryotes. *Microbiological Reviews*, 56(1), pp.100-122.
- Bradham, L.S., 1977.** Fluoride activation of rat brain adenylate cyclase: the requirement for a protein co-factor. *Journal of Cyclic Nucleotide Research*, 3(2), pp.119-128.
- Braun, T. and Dods, R.F., 1975.** Development of a Mn²⁺ sensitive, "soluble" adenylate cyclase in rat testis. *Proceedings of the National Academy of Sciences*, 72(3), pp.1097-1101.
- Bray, E.A., 2000.** Response to abiotic stress. *Biochemistry and Molecular Biology of Plants*, pp.1158-1203.
- Brevedan, R.E. and Egli, D.B., 2003.** Short periods of water stress during seed filling, leaf senescence, and yield of soybean. *Crop Science*, 43(6), pp.2083-2088.
- Brownstein, M.J., Carpten, J.D. and Smith, J.R., 1996.** Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques*, 20(6), pp.1004-1010.
- Cahoon, E.B., 2003.** Genetic Enhancement of Soybean Oil for Industrial Uses: Prospects and Challenges. *Agricultural Biotechnology Forum*, 6(1&2), pp.11-13.
- Cao, D., Li, Y., Liu, B., Kong, F. and Tran, L.S.P., 2018.** Adaptive mechanisms of soybean grown on salt-affected soils. *Land Degradation & Development*, 29(4), pp.1054-1064.
- Carricarte, V.C., Bianchini, G.M., Muschietti, J.P., Tellez-Inon, M.T., Peticari, A., Torres, N. and Flawia, M.M., 1988.** Adenylate cyclase activity in a higher plant, *Alfalfa (Medicago sativa)*. *Biochemical Journal*, 249(3), pp.807-811.
- Castillo-Davis, C.I., Mekhedov, S.L., Hartl, D.L., Koonin, E.V. and Kondrashov, F.A., 2002.** Selection for short introns in highly expressed genes. *Nature Genetics*, 31(4), pp.415-418.
- Chatukuta, P., Dikobe, T.B., Kawadza, D.T., Sehlabane, K.S., Takundwa, M.M., Wong, A., Gehring, C. and Ruzvidzo, O., 2018.** An Arabidopsis clathrin assembly protein with a

predicted role in plant defense can function as an adenylate cyclase. *Biomolecules*, 8(2), p.15.

Campos, M.A., Silva, M.S., Magalhães, C.P., Ribeiro, S.G., Sarto, R.P., Vieira, E.A. and Grossi de Sá, M.F., 2008. Expression in *Escherichia coli*, purification, refolding and antifungal activity of an osmotin from *Solanum nigrum*. *Microbial Cell Factories*, 7(1), pp.1-10.

Chai, C., Wang, Y., Joshi, T., Valliyodan, B., Prince, S., Michel, L., Xu, D. and Nguyen, H.T., 2015. Soybean transcription factor ORFeome associated with drought resistance: a valuable resource to accelerate research on abiotic stress resistance. *BMC genomics*, 16(1), pp.1-9.

Chen, Y., Cann, M.J., Litvin, T.N., Iourgenko, V., Sinclair, M.L., Levin, L.R. and Buck, J., 2000. Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor. *Science*, 289(5479), pp.625-628.

Chen, B.Y., Janes, H.W. and Chen, S., 2002. Computer programs for PCR primer design and analysis. *PCR Cloning Protocols*, 192, p.19.

Chen, J., Xu, W., Velten, J., Xin, Z. and Stout, J., 2012. Characterization of maize inbred lines for drought and heat tolerance. *Journal of Soil and Water Conservation*, 67(5), pp.354-364.

Conti, M. and Beavo, J., 2007. Biochemistry and physiology of cyclic nucleotide phosphodiesterases: essential components in cyclic nucleotide signaling. *Annual Review of Biochemistry*, 76, pp.481-511.

Cooke, C.J., Smith, C.J., Walton, T.J. and Newton, R.P., 1994. Evidence that cyclic AMP is involved in the hypersensitive response of *Medicago sativa* to a fungal elicitor. *Phytochemistry*, 35(4), pp.889-895.

Côté, R., Reisinger F., Martens, L., Barsnes, H., Vizcaino, J.A. and Hermjakob, H., 2010. The ontology lookup service: Bigger and better. *Nucleic Acids Research*, 38(2), pp.155-160.

- Cotta, M.A., Whitehead, T.R. and Wheeler, M.B., 1998.** Identification of a novel adenylate cyclase in the *Ruminal anaerobe*, *Prevotella ruminicola* D31d. *FEMS Microbiology Letters*, 164(2), pp.257-260.
- Cutforth, H.W., McGinn, S.M., McPhee, K.E. and Miller, P.R., 2007.** Adaptation of pulse crops to the changing climate of the Northern Great Plains. *Agronomy Journal*, 99(6), pp.1684-1699.
- Danchin, A., 1993.** Phylogeny of adenyl cyclases. *Advances in second messenger and Phosphoprotein Research*, 27, pp.109-109.
- Das, S. and Bansal, M., 2019.** Variation of gene expression in plants is influenced by gene architecture and structural properties of promoters. *PLoS One*, 14(3), p.e0212678.
- Dlamini, T.S., Tshabalala, P. and Mutengwa, T., 2014.** Soybeans production in South Africa. *OCL - Oilseeds and Fats, Crops and Lipids*, 21(2).
- Dmitriev, A.P., 2003.** Signal molecules for plant defense responses to biotic stress. *Russian Journal of Plant Physiology*, 50(3), pp.417-425.
- Dole, S., Klingen, Y., Nagarajavel, V. and Schnetz, K., 2004.** The protease Lon and the RNA-binding protein Hfq reduce silencing of the *Escherichia coli* *bgl* operon by H-NS. *Journal of Bacteriology*, 186(9), pp.2708-2716.
- Ehsan, H., Reichheld, J.P., Roef, L., Witters, E., Lardon, F., Van Bockstaele, D., Van Montagu, M., Inzé, D. and Van Onckelen, H., 1998.** Effect of indomethacin on cell cycle dependent cyclic AMP fluxes in tobacco BY-2 cells. *FEBS Letters*, 422(2), pp.165-169.
- Faustino, F.C., Garcia, R.N., Agtarap, M.L., Tecson-Mendoza, E.M. and Lips, S.H., 2000.** Salt tolerance in corn: growth responses, ion accumulation, nitrate reductase and PEP-carboxylase activities. *Philippine Journal of Crop Science*, 25(1), pp.17-26.
- Ferrarini, A., De Stefano, M., Baudouin, E., Pucciariello, C., Polverari, A., Puppo, A. and Delledonne, M., 2008.** Expression of *Medicago truncatula* genes responsive to nitric oxide

in pathogenic and symbiotic conditions. *Molecular plant-microbe interactions*, 21(6), pp.781-790.

Fido, R.J., Mills, E.C., Rigby, N.M. and Shewry, P.R., 2004. Protein extraction from plant tissues. *Protein Purification Protocols*, 244, pp.21-27.

Gage, D.J., 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiology and molecular biology reviews*, 68(2), pp.280-300.

Gao, C., Jiang, B., Wang, Y., Liu, G. and Yang, C., 2012. Overexpression of a heat shock protein (ThHSP18. 3) from *Tamarix hispida* confers stress tolerance to yeast. *Molecular Biology Reports*, 39(4), pp.4889-4897.

García, J.E., Maroniche, G., Creus, C., Suárez-Rodríguez, R., Ramirez-Trujillo, J.A. and Groppa, M.D., 2017. *In vitro* PGPR properties and osmotic tolerance of different *Azospirillum* native strains and their effects on growth of maize under drought stress. *Microbiological Research*, 202, pp.21-29.

Gehring, C., 2010. Adenyl cyclases and cAMP in plant signaling-past and present. *Cell Communication and Signaling*, 8(1), pp.1-5.

Gehring, C. and Turek, I.S., 2017. Cyclic nucleotide monophosphates and their cyclases in plant signaling. *Frontiers in Plant Science*, 8, pp.1704.

Gottig, N., Garavaglia, B.S., Daurelio, L.D., Valentine, A., Gehring, C., Orellano, E.G. and Ottado, J., 2009. Modulating host homeostasis as a strategy in the plant-pathogen arms race. *Communicative & Integrative Biology*, 2(2), pp.89-90.

Gouda, S., Kerry, R.G., Das, G., Paramithiotis, S., Shin, H.S. and Patra, J.K., 2018. Revitalization of plant growth promoting rhizobacteria for sustainable development in agriculture. *Microbiological Research*, 206, pp.131-140.

- Gross, M.K., Au, D.C., Smith, A.L. and Storm, D.R., 1992.** Targeted mutations that ablate either the adenylate cyclase or hemolysin function of the bifunctional *cyaA* toxin of *Bordetella pertussis* abolish virulence. *Proceedings of the National Academy of Sciences*, 89(11), pp.4898-4902.
- Grossman, T.H., Kawasaki, E.S., Punreddy, S.R. and Osburne, M.S., 1998.** Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene*, 209(1-2), pp.95-103.
- Gupta, B. and Huang, B., 2014.** Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization. *International Journal of Genomics*, 2014, p.701596.
- Gupta, S.K. and Shukla, P., 2016.** Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. *Critical Reviews in Biotechnology*, 36(6), pp.1089-1098.
- Han, M., Wu, W., Wu, W.H. and Wang, Y., 2016.** Potassium transporter KUP7 is involved in K⁺ acquisition and translocation in Arabidopsis root under K⁺-limited conditions. *Molecular Plant*, 9(3), pp.437-446.
- Hall, R.A., De Sordi, L., MacCallum, D.M., Topal, H., Eaton, R., Bloor, J.W., Robinson, G.K., Levin, L.R., Buck, J., Wang, Y. and Gow, N.A., 2010.** CO₂ acts as a signalling molecule in populations of the fungal pathogen *Candida albicans*. *PLoS Pathogens*, 6(11), p.e1001193.
- Hanoune, J. and Defer, N., 2001.** Regulation and role of adenylyl cyclase isoforms. *Annual Review of Pharmacology and Toxicology*, 41(1), pp.145-174.
- Holzinger, A., Phillips, K.S. and Weaver, T.E., 1996.** Single-step purification/solubilization of recombinant proteins: application to surfactant protein B. *Biotechniques*, 20(5), pp.804-808.

- Irving, H.R., Cahill, D.M. and Gehring, C., 2018.** Moonlighting proteins and their role in the control of signaling microenvironments, as exemplified by cGMP and phytosulfokine receptor 1 (PSKR1). *Frontiers in Plant Science*, 9, p.415.
- Isner, J.C. and Maathuis, F.J., 2011.** Measurement of cellular cGMP in plant cells and tissues using the endogenous fluorescent reporter FlincG. *The Plant Journal*, 65(2), pp.329-334.
- Isner, J.C. and Maathuis, F.J., 2016.** cGMP signalling in plants: from enigma to mainstream. *Functional Plant Biology*, 45(2), pp.93-101.
- Ito, M., Takahashi, H., Sawasaki, T., Ohnishi, K., Hikichi, Y. and Kiba, A., 2014.** 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*, 9(1), p.e27420.
- Jha, S.K, Sharma, M. and K Pandey, G., 2016.** Role of cyclic nucleotide gated channels in stress management in plants. *Current Genomics*, 17(4), pp.315-329.
- Jiang, J., Fan, L.W. and Wu, W.H., 2005.** Evidence for involvement of endogenous cAMP in Arabidopsis defense responses to *Verticillium* toxins. *Cell Research*, 15(8), pp.585-592.
- Jin, X.C. and Wu, W.H., 1999.** Involvement of cyclic AMP in ABA-and Ca^{2+} -mediated signal transduction of stomatal regulation in *Vicia faba*. *Plant and Cell Physiology*, 40(11), pp.1127-1133.
- Kasahara, M., Suetsugu, N., Urano, Y., Yamamoto, C., Ohmori, M., Takada, Y., Okuda, S., Nishiyama, T., Sakayama, H., Kohchi, T. and Takahashi, F., 2016.** An adenylyl cyclase with a phosphodiesterase domain in basal plants with a motile sperm system. *Scientific Reports*, 6(1), pp.1-11.
- Kacperska, A., 2004.** Sensor types in signal transduction pathways in plant cells responding to abiotic stressors: do they depend on stress intensity?. *Physiologia Plantarum*, 122(2), pp.159-168.

- Khaitovich, P., Weiss, G., Lachmann, M., Hellmann, I., Enard, W., Muetzel, B., Wirkner, U., Ansorge, W., Pääbo, S. and Botstein, D., 2004.** A neutral model of transcriptome evolution. *PLoS Biology*, 2(5), p.e132.
- Kim, I.S., Kim, C.H. and Yang, W.S., 2021.** Physiologically active molecules and functional properties of soybeans in human health - A current perspective. *International Journal of Molecular Sciences*, 22(8), p.4054.
- Kim, I.S., Hwang, C.W., Yang, W.S. and Kim, C.H., 2021.** Current perspectives on the physiological activities of fermented soybean-derived cheonggukjang. *International Journal of Molecular Sciences*, 22(11), p.5746.
- Kneusel, R.E., Crowe, J., Wulbeck, M. and Ribbe, J., 1998.** Procedures for the analysis and purification of His-tagged proteins. *Methods in Molecular Medicine*, 13, pp. 293–308.
- Künstler, A., Bacsó, R., Gullner, G., Hafez, Y.M. and Király, L., 2016.** Staying alive—is cell death dispensable for plant disease resistance during the hypersensitive response?. *Physiological and Molecular Plant Pathology*, 93, pp.75-84.
- Kudla, J., Becker, D., Grill, E., Hedrich, R., Hippler, M., Kummer, U., Parniske, M., Romeis, T. and Schumacher, K., 2018.** Advances and current challenges in calcium signaling. *New Phytologist*, 218(2), pp.414-431.
- Kunert, K.J., Vorster, B.J., Fenta, B.A., Kibido, T., Dionisio, G. and Foyer, C.H., 2016.** Drought stress responses in soybean roots and nodules. *Frontiers in Plant Science*, 7, p.1015.
- Kuo, J.T., Chang, Y.J. and Tseng, C.P., 2003.** Growth rate regulation of lac operon expression in *Escherichia coli* is cyclic AMP dependent. *FEBS Letters*, 553(3), pp.397-402.
- Kusnadi, A.R., Nikolov, Z.L. and Howard, J.A., 1997.** Production of recombinant proteins in transgenic plants: practical considerations. *Biotechnology and Bioengineering*, 56(5), pp.473-484.

- Kwezi, L., Ruzvidzo, O., Wheeler, J.I., Govender, K., Iacuone, S., Thompson, P.E., Gehring, C. and Irving, H.R., 2011.** The phytosulfokine (PSK) receptor is capable of guanylate cyclase activity and enabling cyclic GMP-dependent signaling in plants. *Journal of Biological Chemistry*, 286(25), pp.22580-22588.
- Latef, A.A.H.A. and Ahmad, P., 2015.** Legumes and breeding under abiotic stress: an overview. *Legumes under environmental stress: Yield, improvement and adaptations*, pp.1-20.
- Laranjo, M., Alexandre, A. and Oliveira, S., 2014.** Legume growth-promoting rhizobia: an overview on the *Mesorhizobium* genus. *Microbiological research*, 169(1), pp.2-17.
- Li, X., Borsics, T., Harrington, H.M. and Christopher, D.A., 2005.** Arabidopsis AtCNGC10 rescues potassium channel mutants of *E. coli*, yeast and Arabidopsis and is regulated by calcium/calmodulin and cyclic GMP in *E. coli*. *Functional Plant Biology*, 32(7), pp.643-653.
- Liang, S., Sun, J., Luo, Y., Lv, S., Chen, J., Liu, Y. and Hu, X., 2022.** cAMP is a promising regulatory molecule for plant adaptation to heat stress. *Life*, 12(6), p.885.
- Lindwall, G., Chau, M.F., Gardner, S.R. and Kohlstaedt, L.A., 2000.** A sparse matrix approach to the solubilization of overexpressed proteins. *Protein Engineering*, 13(1), pp.67-71.
- Liu, F., Andersen, M.N. and Jensen, C.R., 2003.** Loss of pod set caused by drought stress is associated with water status and ABA content of reproductive structures in soybean. *Functional Plant Biology*, 30(3), pp.271-280.
- Maathuis, F.J. and Sanders, D., 2001.** Sodium uptake in Arabidopsis roots is regulated by cyclic nucleotides. *Plant Physiology*, 127(4), pp.1617-1625.
- MacConkey, A., 1905.** Lactose-fermenting bacteria in faeces. *Journal of Epidemiology and Infection*, 5(3):333-379.
- Maitra, S., 2018.** SOYBEAN (BHATMAN). In *Forage Crops of the World, Volume II: Minor Forage Crops*, pp. 121-140.

- Marondedze, C., Wong, A., Thomas, L., Irving, H. and Gehring, C., 2015.** Cyclic nucleotide monophosphates in plants and plant signaling. *Non-canonical Cyclic Nucleotides*, pp.87-103.
- Masuda, T. and Goldsmith, P.D., 2009.** World soybean production: area harvested, yield, and long-term projections. *International food and agribusiness management review*, 12(1030-2016-82753), pp.1-20.
- Mihic, S.J and Harris, R.A., 1996.** Single step purification/solubilisation of recombinant proteins: Application to surfactant protein B. *Biotechniques*, 20, pp 804-808
- Miller, G., Schlauch, K., Tam, R., Cortes, D., Torres, M.A., Shulaev, V., Dangi, J.L. and Mittler, R., 2009.** The plant NADPH oxidase RBOHD mediates rapid systemic signaling in response to diverse stimuli. *Science Signaling*, 2(84), p.45.
- Moustafa-Farag, M., Almoneafy, A., Mahmoud, A., Elkelish, A., Arnao, M.B., Li, L. and Ai, S., 2019.** Melatonin and its protective role against biotic stress impacts on plants. *Biomolecules*, 10(1), p.54.
- Moutinho, A., Hussey, P.J., Trewavas, A.J. and Malho, R., 2001.** cAMP acts as a second messenger in pollen tube growth and reorientation. *Proceedings of the National Academy of Sciences*, 98(18), pp.10481-10486.
- Nawaz, Z., Kakar, K.U., Saand, M.A. and Shu, Q.Y., 2014.** Cyclic nucleotide-gated ion channel gene family in rice, identification, characterization and experimental analysis of expression response to plant hormones, biotic and abiotic stresses. *Bmc Genomics*, 15(1), pp.1-18.
- Ohmori, M., Ohmori, K. and Hasunuma, K., 1988.** Rapid change in cyclic 3', 5'-AMP concentration triggered by a light-off or light-on signal in *Anabaena cylindrica*. *Archives of Microbiology*, 150(2), pp.203-204.
- Ohmori, M., Terauchi, K., Okamoto, S. and Watanabe, M., 2002.** Regulation of cAMP-mediated Photosignaling by a Phytochrome in the Cyanobacterium *Anabaena cylindrica*. *Photochemistry and photobiology*, 75(6), pp.675-679.

- Pacini, B., Petrigliano, A., Diffley, P., Paffetn, A., Brown, E.G., Martelli, P., Trabalzini, L., Bovalini, L., Lhusini, P. and Newton, R.P., 1993.** Adenylyl cyclase activity in roots of *Pisum sativum*. *Phytochemistry*, 34(4), pp.899-903.
- Pastan, I.R.A. and Adhya, S., 1976.** Cyclic adenosine 5'-monophosphate in *Escherichia coli*. *Bacteriological Reviews*, 40(3), pp.527-551.
- Perlman, R.L. and Pastan, I., 1969.** Pleiotropic deficiency of carbohydrate utilization in an adenyl cyclase deficient mutant of *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 37(1), pp.151-157.
- Pandey, G.K. and Sanyal, S.K., 2020.** *Functional dissection of calcium homeostasis and transport machinery in plants*. Springer Nature.
- Phang, T.H., Shao, G. and Lam, H.M., 2008.** Salt tolerance in soybean. *Journal of Integrative Plant Biology*, 50(10), pp.1196-1212.
- Pizzino, G., Irrera, N., Cucinotta, M., Pallio, G., Mannino, F., Arcoraci, V., Squadrito, F., Altavilla, D. and Bitto, A., 2017.** Oxidative Stress: Harms and Benefits for Human Health. *Oxidative medicine and cellular longevity*, 2017, p.8416763.
- Rato, C., Monteiro, D., Hepler, P.K. and Malhó, R., 2004.** Calmodulin activity and cAMP signalling modulate growth and apical secretion in pollen tubes. *The Plant Journal*, 38(6), pp.887-897.
- Ray, N.G., Vieth, W.R. and Venkatasubramanian, K., 1987.** Regulation of *lac* operon expression in mixed sugar chemostat cultures. *Biotechnology and Bioengineering*, 29(8), pp.1003-1014.
- Ren, X.Y., Vorst, O., Fiers, M.W., Stiekema, W.J. and Nap, J.P., 2006.** In plants, highly expressed genes are the least compact. *Trends in Genetics*, 22(10), pp.528-532.
- Richards, H., Das, S., Smith, C.J., Pereira, L., Geisbrecht, A., Devitt, N.J., Games, D.E., van Geyschem, J., Brenton, A.G. and Newton, R.P., 2002.** Cyclic nucleotide content of tobacco BY-2 cells. *Phytochemistry*, 61(5), pp.531-537.

- Rosano, G.L. and Ceccarelli, E.A., 2014.** Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5, p.172.
- Ruzvidzo, O., Dikobe, B.T., Kawadza, D.T., Mabadahanye, G.H., Chatukuta, P. and Kwezi, L., 2013.** Recombinant expression and functional testing of candidate adenylate cyclase domains. *Cyclic Nucleotide Signaling in Plants*, pp.13-25.
- Saed-Moucheshi, A., Pakniyat, H., Pirasteh-Anosheh, H. and Azooz, M.M., 2014.** Role of ROS as signaling molecules in plants. *Oxidative Damage to Plants*, pp.585-620.
- Saeed, M.T., Wahid, M.A., Saleem, M.F., Cheema, M.A. and Shahid, M., 2017.** Improving the stand establishment, phenology and yield of soybean (*Glycine max* L.) by various physiological enhancements. *Pakistan Journal of Agricultural Research*, 30(3), pp.218-225.
- Sakai, T. and Kogiso, M., 2008.** Soy isoflavones and immunity. *The Journal of Medical Investigation*, 55(3 & 4), pp.167-173.
- Schenk, P.W. and Snaar-Jagalska, B.E., 1999.** Signal perception and transduction: the role of protein kinases. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1449(1), pp.1-24.
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J. and Xu, D., 2010.** Genome sequence of the palaeopolyploid soybean. *Nature*, 463(7278), pp.178-183.
- Serrano, R., Mulet, J.M., Rios, G., Marquez, J.A., De Larrinoa, I.F., Leube, M.P., Mendizabal, I., Pascual-Ahuir, A., Proft, M., Ros, R. and Montesinos, C., 1999.** A glimpse of the mechanisms of ion homeostasis during salt stress. *Journal of experimental botany*, pp.1023-1036.
- Shah, S. and Peterkofsky, A., 1991.** Characterization and generation of *Escherichia coli* adenylate cyclase deletion mutants. *Journal of Bacteriology*, 173(10), pp.3238-3242.

- Sharma, P., Jha, A.B., Dubey, R.S. and Pessarakli, M., 2012.** Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions. *Journal of Botany*, 2012, pp.34-59.
- Shuman, S., 1992.** Two classes of DNA end-joining reactions catalyzed by vaccinia topoisomerase I. *Journal of Biological Chemistry*, 267(24), pp.16755-16758.
- Sorroche, F., Walch, M., Zou, L., Rengel, D., Maillet, F., Gibelin-Viala, C., Poinso, V., Chervin, C., Masson-Boivin, C., Gough, C. and Batut, J., 2019.** Endosymbiotic Sinorhizobium meliloti modulate Medicago root susceptibility to secondary infection via ethylene. *New Phytologist*, 223(3), pp.1505-1515.
- Steer, M.L. and Levitzki, A., 1975.** The control of adenylate cyclase by calcium in turkey erythrocyte ghosts. *Journal of Biological Chemistry*, 250(6), pp.2080-2084.
- Stempfer, G., Höll-Neugebauer, B. and Rudolph, R., 1996.** Improved refolding of an immobilized fusion protein. *Nature Biotechnology*, 14(3), pp.329-334.
- Stühmer, W., Conti, F., Suzuki, H., Wang, X., Noda, M., Yahagi, N., Kubo, H. and Numa, S., 1989.** Structural parts involved in activation and inactivation of the sodium channel. *Nature*, 339(6226), pp.597-603.
- Świeżawska, B., Duszyn, M., Jaworski, K. and Szmidt-Jaworska, A., 2018.** Downstream targets of cyclic nucleotides in plants. *Frontiers in Plant Science*, 9, p.1428.
- Świeżawska, B., Jaworski, K., Pawelek, A., Grzegorzewska, W., Szewczuk, P. and Szmidt-Jaworska, A., 2014.** Molecular cloning and characterization of a novel adenylyl cyclase gene, HpAC1, involved in stress signaling in *Hippeastrum hybridum*. *Plant Physiology and Biochemistry*, 80, pp.41-52.
- Takahashi, F. and Shinozaki, K., 2019.** Long-distance signaling in plant stress response. *Current opinion in plant biology*, 47, pp.106-111.

- Tavakkoli, E., Rengasamy, P. and McDonald, G.K., 2010.** High concentrations of Na⁺ and Cl⁻ ions in soil solution have simultaneous detrimental effects on growth of faba bean under salinity stress. *Journal of experimental botany*, 61(15), pp.4449-4459.
- Terakado, J., Okamura, M., Fujihara, S., Ohmori, M. and Yoneyama, T., 1997.** Cyclic AMP in rhizobia and symbiotic nodules. *Annals of Botany*, 80(4), pp.499-503.
- Tian, C.F., Garnerone, A.M., Mathieu-Demazière, C., Masson-Boivin, C. and Batut, J., 2012.** Plant-activated bacterial receptor adenylate cyclases modulate epidermal infection in the *Sinorhizobium meliloti* – Medicago symbiosis. *Proceedings of the National Academy of Sciences*, 109(17), pp.6751-6756.
- Timmers, A.C., Soupène, E., Auriac, M.C., de Billy, F., Vasse, J., Boistard, P. and Truchet, G., 2000.** Saprophytic intracellular rhizobia in alfalfa nodules. *Molecular Plant-Microbe Interactions*, 13(11), pp.1204-1213.
- Thomas, L., Marondedze, C., Ederli, L., Pasqualini, S. and Gehring, C., 2013.** Proteomic signatures implicate cAMP in light and temperature responses in *Arabidopsis thaliana*. *Journal of Proteomics*, 83, pp.47-59.
- Topal, H., Fulcher, N.B., Bitterman, J., Salazar, E., Buck, J., Levin, L.R., Cann, M.J., Wolfgang, M.C. and Steegborn, C., 2012.** Crystal structure and regulation mechanisms of the CyaB adenyl cyclase from the human pathogen *Pseudomonas aeruginosa*. *Journal of Molecular Biology*, 416(2), pp.271-286.
- Tuteja, N., 2007.** Mechanisms of high salinity tolerance in plants. *Methods in Enzymology*, 428, pp.419-438.
- Upchurch, R.G. and Elkan, G.H., 1978.** The role of ammonia, L-glutamate, and cyclic adenosine 3', 5'-monophosphate in the regulation of ammonia assimilation in *Rhizobium japonicum*. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 538(2), pp.244-248.

- Ullmann, A. and Danchin, A., 1983.** Role of cyclic AMP in bacteria. *Advances in Cyclic Nucleic Acids and Respiratory*, 15, pp. 1-53.
- Van Damme, T., Blancquaert, D., Couturon, P., Van Der Straeten, D., Sandra, P. and Lynen, F., 2014.** Wounding stress causes rapid increase in concentration of the naturally occurring 2',3'-isomers of cyclic guanosine-and cyclic adenosine monophosphate (cGMP and cAMP) in plant tissues. *Phytochemistry*, 103, pp.59-66.
- Van Loon, L.C., Rep, M. and Pieterse, C.M., 2006.** Significance of inducible defense-related proteins in infected plants. *Annual Review of Phytopathology*, 44, pp.135-162.
- Valsecchi, F., Konrad, C. and Manfredi, G., 2014.** Role of soluble adenylyl cyclase in mitochondria. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1842(12), pp.2555-2560.
- Volotovskii, I.D., Sokolovsky, S.G., Molchan, O.V. and Knight, M.R., 1998.** Second messengers mediate increases in cytosolic calcium in tobacco protoplasts. *Plant Physiology*, 117(3), pp.1023-1030.
- Wang, W., Vinocur, B. and Altman, A., 2003.** Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta*, 218(1), pp.1-14.
- Witters, E., Roef, L., Newton, R.P., Van Dongen, W., Esmans, E.L. and Van Onckelen, H.A., 1996.** Quantitation of cyclic nucleotides in biological samples by negative electrospray tandem mass spectrometry coupled to ion suppression liquid chromatography. *Rapid Communications in Mass Spectrometry*, 10(2), pp.225-231.
- Yamaguchi, H. and Miyazaki, M., 2014.** Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules*, 4(1), pp.235-251.
- Yang, H., Zhao, Y., Chen, N., Liu, Y., Yang, S., Du, H., Wang, W., Wu, J., Tai, F., Chen, F. and Hu, X., 2021.** A new adenylyl cyclase, putative disease-resistance RPP13-like protein 3,

participates in abscisic acid-mediated resistance to heat stress in maize. *Journal of Experimental Botany*, 72(2), pp.283-301.

Yang, H., You, C., Yang, S., Zhang, Y., Yang, F., Li, X., Chen, N., Luo, Y. and Hu, X., 2021. The role of calcium/calcium-dependent protein kinases signal pathway in pollen tube growth. *Frontiers in Plant Science*, 12, p.346.

Yuan, P., Jauregui, E., Du, L., Tanaka, K. and Poovaiah, B.W., 2017. Calcium signatures and signaling events orchestrate plant–microbe interactions. *Current Opinion in Plant Biology*, 38, pp.173-183.

Yunghans, W.N. and Morré, D.J., 1977. Adenylate cyclase activity not found in soybean hypocotyl and onion meristem. *Plant Physiology*, 60(1), pp.144-149.

Zhang, G., Fang, X., Guo, X., Li, L.I., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H. and Xiong, Z., 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*, 490(7418), pp.49-54.

Zhang, Y., Olsen, D.R., Nguyen, K.B., Olson, P.S., Rhodes, E.T. and Mascarenhas, D., 1998. Expression of eukaryotic proteins in soluble form in *Escherichia coli*. *Protein Expression and Purification*, 12(2), pp.159-165.

Zhao, J., Davis, L.C. and Verpoorte, R., 2005. Elicitor signal transduction leading to production of plant secondary metabolites. *Biotechnology advances*, 23(4), pp.283-333.

Zhao, Y., Liu, Y., Ji, X., Sun, J., Lv, S., Yang, H., Zhao, X. and Hu, X., 2021. Physiological and proteomic analyses reveal cAMP-regulated key factors in maize root tolerance to heat stress. *Food and Energy Security*, 10(4), p.e309.

Zhou, M.Y. and Gomez-Sanchez, C.E., 2000. Universal TA cloning. *Current Issues in Molecular Biology*, 2(1), pp.1-7.