Communications in PLANT SCIENCES

http://cpsjournal.org
e-ISSN 2237-4027
Volume 12 (2022)

RESEARCH ARTICLE

Identification and characterization of a soybean protein with adenylyl cyclase activity

Enetia Desberia Bobo, **Katlego Selaelo Sehlabane**, **Tshegofatso Bridget Dikobe**, **Mutsa Monica Takundwa**, **David Tonderai Kawadza** and **Oziniel Ruzvidzo***

North-West University, Mmabatho, North West, South Africa.
Correspondence to: Oziniel.Ruzvidzo@nwu.ac.za

Submitted on May 05, 2022
Accepted on November 14, 2022
Early View on November 14, 2022
Final Publication on November 23, 2022

Authors declare no conflict of interest

KEYWORDS:
Glycine max
GmAC
cAMP
Co-expression
Stress tolerance
Early plant development

ABSTRACT

Soybean [*Glycine max* (L.) Merrill] is a high value leguminous crop characterized by its excellent protein content and ability to improve soil quality through nitrogen fixation. Whereas this plant has attractive human and animal feed attributes in addition to its pharmaceutical and industrial uses, its growth and yield are severely affected by drought. Thus any research aimed at understanding the genome response of this plant to drought and other related environmental stress factors would be worthwhile. In plants, in general, second messengers have a key role in linking and coordinating environmental stimuli to cellular communication and responses. One group of such messengers are adenylyl cyclases (ACs) and their catalytic product 3′,5′-cyclic adenosine monophosphate (cAMP), involved in plant growth, cell division, reproduction, development and response to stress. However, while ACs have been reported in some plant species such as Arabidopsis and maize, their presence together with their cAMP-dependent systems in *G. max* have largely remained unavailable. Fortunately, a putative molecule, Glyma.07G251000 (accession number: XP_003529590), with a predicted function as an AC in *G. max* has at some point been reported. This molecule harbors a domain annotated AC catalytic center and therefore, was herein targeted for study. In order to characterize the Glyma.07G251000, we cloned and expressed it, followed by purification of the resultant recombinant protein (GmAC). When tested *in vitro* for AC activity, the GmAC protein showed a Mn^{2+}-dependent activity that is positively enhanced by calcium. GmAC also complemented the AC-deficiency (*cya*A mutation) of an SP850 mutant strain when expressed in *Escherichia coli*. When analysed by a web-based approach system, the GmAC protein was found to be co-expressed and co-regulated with various other proteins responsible for early plant development and stress response, strongly suggesting that it has a central role in these two key cellular processes. In addition, the GmAC protein conferred stress resistance to EXPRESS BL21 (DE3) pLysS DUOs cells when expressed in these host cells under salt (200 mM NaCl) and oxidative stress (0.2 mM H_2O_2). Conceivably, our findings showed that GmAC is an AC protein with a role in early plant development and stress response.

Highlighted Conclusions

1. GmAC is an adenylyl cyclase and the first ever such protein to be identified in soybean.
2. GmAC confers stress tolerance to *Escherichia coli* and is co-expressed/co-regulated with other soybean proteins responsible for early plant development and stress response.

INTRODUCTION

Soybean or *Glycine max* is one of the most important legume crops that provide excellent source of oil and quality proteins for livestock and humankind. Apart from being consumable, soybean products have been gaining attention for their other additional attributes such as the anti-cancer properties in pharmaceuticals (Ko et al. 2013) and the protein-based bio-degradable properties for possible consideration as alternatives in the plastic industry (Song et al. 2011). These diverse attributes of soybean make this legume a more widely desired crop plant, whose demand is rapidly increasing. Nevertheless, soybean production is severely affected by drought (Brevedan and Egli 2003, Liu et al. 2003) and other related extreme conditions such as cold, floods, heat, pests and pathogens (Deshmukh et al. 2014). Notably, plants in general, respond and adapt to various biotic and abiotic environmental stress factors with an array of molecular systems and mechanisms, which among them include the second messenger 3',5'-cyclic

adenosine monophosphate (cAMP) and its generating enzyme adenylyl cyclase (AC). In general, second messengers play a key role in linking environmental stimuli to cellular responses (Frezza et al. 2018) and some of those processes commonly coordinated by cAMP include growth, cell division, reproduction, development and response to stress (Choi and Xu 2010).

To this day, a number of AC candidate molecules together with their cAMP-dependent processes have been reported in some plants but surprisingly, not soybean (Yunghans and Moore 1977). Such candidate molecules include two in *Zea mays*, six in *Arabidopsis thaliana*, and one each in *Nicotiana benthamiana*, *Hippeastrum hybridum* and *Marchantia polymorpha*. The two *Z. mays* ACs are a pollen signaling protein (ZmPSiP) with a role in polarized pollen tube growth and re-orientation (Moutinho et al. 2001) and a putative disease-resistance RPP13-like protein 3 (ZmRPP13-LK3) that participates in abscisic acid (ABA) mediated resistance to heat stress (Yang et al. 2021). The Arabidopsis candidates are a pentatricopeptide repeat protein (AtPPR-AC) annotated for chloroplast biogenesis and restoration of the cytoplasmic male sterility (CMS) (Ruzvidzo et al. 2013), two permeases (AtKUP7 and AtKUP5) responsible for K^+ ion uptake (Al-Younis et al. 2015, 2018), a clathrin assembly protein (AtClAP) that has a predicted role in endocytosis and plant defense (Chatukuta et al. 2018), a leucine-rich repeat protein (AtLRRAC1) responsible for pathogen defense (Bianchet et al. 2019, Ruzvidzo et al. 2019) and a 9-*cis*-epoxycarotenoid dioxygenase (AtNCED3) that has a role in biosynthesis of the stress hormone abscisic acid (ABA) (Al-Younis et al. 2021). The other three are a *N*. *benthamiana* AC protein (NbAC) responsible for the tabtoxinine-β-lactam-induced cell deaths during wildfire diseases (Ito et al. 2014), a *H*. *hybridum* AC protein (HpAC1) involved in stress signalling (Świeżawska et al. 2014), and a *M*. *polymorpha* AC protein (MpCAPE) that has a role in cell and male organ development (Kasahara et al. 2016).

Typically, soybean has a genome size of almost 1.1 gigabases, with approximately 46 430 protein coding genes (Turner et al. 2012). Since its sequencing in 2010 (Schmutz et al. 2010), the majority of its protein coding genes remain experimentally unconfirmed (Chai et al. 2015) and as such, genes responsible for the coding of ACs and/or other related protein molecules could be present in this legume. In line with this, some recent and independent phylogenetic studies on ACs in higher plants, pointed to a possible existence of such molecules in *G*. *max*; Glyma.07G251000 (accession number: XP_003529590) (Ito et al. 2014) and Glyma.15G090100 (accession number: XP_003547191) (Świeżawska et al. 2014). Interestingly, none of these two molecules has yet been experimentally tested and/or functionally characterized. In this study therefore, we focused onto Glyma.07G251000, specifically because it possesses the annotated AC catalytic center, previously proposed by Gehring (2010) and later confirmed in most functional plant ACs (Moutinho et al. 2001, Ruzvidzo et al. 2013, Ito et al. 2014, Al-Younis et al. 2015, 2018, 2021, Chatukuta et al. 2018, Bianchet et al. 2019, Ruzvidzo et al. 2019, Yang et al. 2021). Thus in order to characterize the Glyma.07G251000, we cloned and expressed it, followed by elucidation of the functional activities of its resultant expression product (GmAC).

MATERIAL AND METHODS

Isolation of the Glyma.07G251000 gene and generation of the recombinant GmAC protein. Total RNA was extracted from 2-week-old *G*. *max* seedlings using the RNeasy plant mini kit, in combination with DNase 1 treatment, as instructed by the manufacturer (Qiagen, Crawley, UK). Copy DNA (cDNA) sequence of Glyma.07G251000 was retrieved from UniprotKB (https://www.uniprot.org/) in FASTA (canonical) format and verified for presence of the AC catalytic center using the PROSITE database located within the Expert Protein Analysis System (ExPASy) proteomics server (https://www.expasy.ch/). Glyma.07G251000 cDNA synthesis from the total RNA and its subsequent amplification was performed in the presence of two sequence-specific primers:

forward: 5'-GAAATAGAAAAGTTGAGATCCAATTCAATG-3'

reverse: 5'-GTGTGGAATATCTTCCTTTCTGTCCAGGTC-3'

using a Verso 1-Step RT-PCR kit and in accordance with the manufacturer's instructions (Thermo Scientific, Rockford, USA). The PCR product was then cloned into a pTrcHis2-TOPO expression vector via the TA cloning system (Invitrogen Corp., Carlsbad, USA) to make a pTrcHis2-TOPO:GmAC fusion expression construct with a C-terminus His purification tag. Expression, purification and refolding processes of the recombinant GmAC protein were undertaken as is detailed elsewhere (Kwezi et al. 2011, Chatukuta et al. 2018). The relative molecular mass of the generated GmAC protein was estimated using the ProtParam tool on the ExPasy Proteomics Server (http://au.expasy.org/tool/.protpatram.html). Whereas the purified protein was used for *in vitro* activity assays, the produced pTrcHis2-TOPO:GmAC fusion expression construct was used for complementation testing and stress response activity assays.

Determination of the *in vitro* AC activity of GmAC. The AC activity of the purified recombinant GmAC was assessed *in vitro* by incubating 5 µg of the protein in 50 mM Tris-Cl; pH 8.0, 2 mM isobutylmethylxanthine (IBMX,

Sigma-Aldrich Corp., Missouri, USA) to inhibit phosphodiesterases, 5 mM Mg^{2+} or Mn^{2+} and 1 mM ATP with or without 250 µM Ca^{2+}, in a final volume of 200 µl, followed by measurement of the generated cAMP. Levels of the generated cAMP were determined by enzyme immunoassaying following its acetylation protocol as described by the supplier's manual (Sigma-Aldrich Corp., Missouri, USA; code: CA201). The method is detailed elsewhere (Al-Younis et al. 2015, Chatukuta et al. 2018).

Testing for the ability of GmAC to complement *cya*A mutation in an *E. coli* AC-deficient strain. The *E. coli cya*A mutant SP850 strain (lam-, el4-, relA1, spoT1, *cya*A1400 (:kan), thi-1) (Shah and Peterkofsky 1991), deficient in the adenylyl cyclase (*cya*A) gene, was obtained from the *E. coli* Genetic Stock Centre (Yale University, New Haven, USA) (accession No. 7200). The strain was prepared to be chemically competent followed by its transformation with the pTrcHis2-TOPO:GmAC fusion construct (through heat shock at 42°C for 2 min). The transformed bacteria were then grown at 37°C in Luria-Bertani (LB) media containing ampicillin (100 µg mL^{-1}) and kanamycin (15 µg mL^{-1}) until their cell density had reached an optical density (OD_{600}) of 0.5. The cells were treated with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma-Aldrich Corp., Missouri, USA), for transgene induction, and further incubated for 4 hrs prior to streaking on MacConkey agar. The method is detailed elsewhere (Moutinho et al. 2001, Ruzvidzo et al. 2013, Świeżawska et al. 2014).

Co-expression and stimulus-specific analysis of Glyma.07G251000. The similarity search tool, Co-expression from the Compendium Wide Analysis; Genevestigator v3 (http://www.genevestigator.ethz.ch) (Hruz et al. 2008) was used to assess genes in soybean that are functionally co-expressed with Glyma.07G251000. The search was performed on all available experiments across the [Soybean] Affymetrix Soybean Genome Array platform or GM_AFFY_SOYBEAN database, leaving the gene list limit blank to obtain a full correlational list. From this search, 50 top most co-expressed genes (CEG-50) were considered, based on the Pearson correlation coefficient as a measure of similarity between them. FFPred 2.0, accessed at the interface of the PSIPRED Protein Analysis Workbench (http://bioinf.cs.ucl.ac/psipred) (Minneci et al. 2013), was used to perform stimulus-specific analysis and gene ontology (GO) predictions of the Glyma.07G251000 and its CEG-50 in the soybean plant. The covered ontologies included biological process, molecular function and cellular component. An expression heat map of 20 randomly selected CEGs was then generated using the Multiple Array Viewer program from the Multi-Experiment Viewer (MeV) software package (version 4.2.01) created by The Institute for Genomic Research (TIGR) (Saeed et al. 2003).

Testing for the ability of GmAC to confer stress tolerance to *E. coli*. Liquid and spot culture assays were carried out. For liquid culture assaying, 50 mL of double strength yeast-tryptone (2YT) broth media containing 100 µg mL^{-1} ampicillin and 34 µg mL^{-1} chloramphenicol, and supplemented with either 200 mM NaCl (for salt stress) or 0.2 mM H_2O_2 (for oxidative stress) or no chemical supplementation (for non-stress conditions) were inoculated with 1 mL of the EXPRESS BL21 (DE3) pLysS DUOs cells, harboring the pTrcHis2-TOPO:GmAC fusion construct. The prepared cultures were incubated at 37°C in an orbital shaker at 200 rpm. At an OD_{600} of 0.5 of the cells, the 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 hrs at 37°C at 200 rpm while at every 2-hr point, 1 mL of each culture was collected and its OD_{600} measured using a Helios Epsilon spectrophotometer (Thermo Scientific Inc., Madison, USA). The recorded readings were then plotted into graphs and analysed comparatively. For spot culture assaying, the same transformed EXPRESS BL21 (DE3) pLysS DUOs cells, at an OD_{600} 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 10^0, 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^{-1} ampicillin and 34 µg mL^{-1} chloramphenicol, and supplemented with either 200 mM NaCl (salt stress) or 0.2 mM H_2O_2 (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 hrs. Thereafter, diameters of cell colonies were measured in mm with a ruler, to ascertain the extent of cell growth and then a comparative analysis of the cell colony sizes.

RESULTS

The identification of plant ACs mostly involved querying protein sequences with the modified guanylyl cyclase (GC) search motif (Ludidi and Gehring 2003) at position 3, changing it from [CTGH] to [DE] (Gehring 2010) (Figure 1A). This substitution is based on previous findings, which indicated that the conversion of GCs into ACs and vice versa could be easily achieved through a single mutation in the amino acid that confers substrate specificity (Tucker et al. 1998, Roelofs et al. 2001). When the Glyma.07G251000 amino acid sequence was queried by the AC motif, a matching hit was detected towards its C-terminal end (amino acids 333 – 347) (Figure 1B). The whole genome

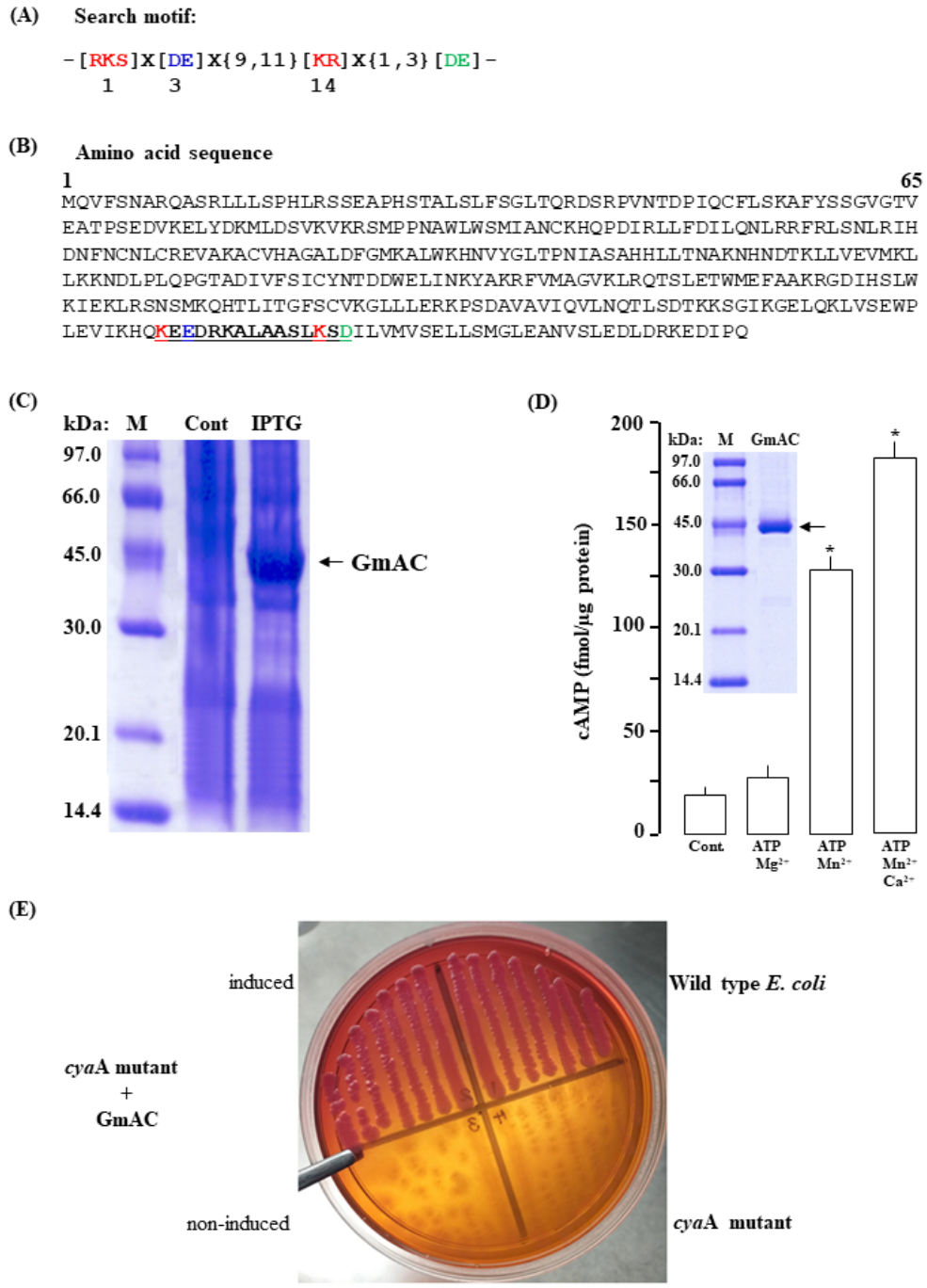

Figure 1. (A) The 14 amino acid AC search motif derived from annotated and experimentally tested GC and AC catalytic centres. The residue forming hydrogen bonding with purine at position 1 is highlighted in red, the residue conferring substrate specificity in position 3 is highlighted in blue, while 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 Mg^{2+}/Mn^{2+}-binding and is colored green. (B) The complete amino acid sequence of GmAC with the AC catalytic center towards its C-terminus (amino acids 333 – 347) highlighted in bold and underline. (C) Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of protein fractions (stained with Coomassie brilliant blue) from the induced (IPTG) and un-induced (Cont) cell cultures, where (M) is the molecular weight marker and the arrow marking the expressed recombinant GmAC protein. (D) cAMP generated by 5 µg recombinant GmAC in the presence of 1 mM ATP and 5 mM Mn^{2+} or Mg^{2+}, or 1 mM ATP and 250 µM Ca^{2+} when 5 mM Mn^{2+} ion is the cofactor. Control reaction contained all other components except the protein and Ca^{2+}. Inset: A Coomassie brilliant blue-stained gel after resolution of the affinity purified His-tagged recombinant GmAC (arrow) by SDS-PAGE. Data are mean values (n = 3) and error bars show standard error (SE) of the mean. Asterisks denote values significantly different from those of control ($p < 0.05$) determined by analysis of variance (ANOVA) and post hoc Student-Newman-Keuls (SNK) multiple range tests. (E) The AC center of GmAC complemented the *cya*A mutant *E. coli* (SP850) to ferment lactose. Wild-type and GmAC-expressing SP850 *E. coli* cells showed a strong reddish colour while both the *cya*A mutant and *cya*A mutant cells with an un-induced recombinant GmAC yielded yellowish colonies.

sequence of Glyma.07G251000 was then cloned into a prokaryotic system and expressed into a 44.97 kDa His-tagged recombinant GmAC protein (Figure 1C). To test if the AC center of GmAC can generate cAMP *in vitro*, the expressed recombinant GmAC was extracted and affinity purified (Figure 1D, inset). The AC activity of the purified recombinant was then tested in a reaction mixture containing ATP as substrate, Mn^{2+} or Mg^{2+} as cofactor, and Ca^{2+} as modulator, followed by measurement of cAMP by enzyme immunoassay. Maximum activity was reached after 20 minutes of the reaction system, generating about 127 fmols μg^{-1} protein of cAMP in the presence of Mn^{2+} and approximately 26 fmols μg^{-1} protein of cAMP in the presence of Mg^{2+} compared to only about 14 fmols μg^{-1} protein of cAMP of the control reaction (Figure 1D). Besides being Mn^{2+}-dependent, the catalytic activity of GmAC was also significantly enhanced by Ca^{2+}, reaching an activity level of around 181 fmols μg^{-1} protein of cAMP when Mn^{2+} is the co-factor (Figure 1D).

In order to investigate if the AC centre of GmAC can rescue an *E. coli* AC-deficient mutant, the Glyma.07G251000 was cloned and expressed in an *E. coli* SP850 strain lacking the AC (*cya*A), essential for lactose fermentation (Ullmann and Danchin 1980, Shah and Peterkofsky 1991, Tang et al. 1995). As a result of the *cya*A mutation, the AC deficient and un-induced transformed *E. coli* cells remained yellowish in colour when grown on MacConkey agar. In contrast, the Glyma.07G251000-transformed SP850 cells, when induced with 0.5 mM IPTG, formed deep reddish colonies much like the wild-type *E. coli* (Figure 1E) thus indicating a functional AC center in the recombinant GmAC protein.

When Glyma.07G251000 was analyzed for co-expression in the *G. max* genome, we noted that 50 of its most correlated genes have high r values of between 0.64 and 0.80 (Appendix A, Table S1). These expression-correlated genes (CEG50) are also significantly enriched for the "biological process" gene ontology (GO) categories 'developmental process' and 'cellular response to stimulus' and "molecular function" GO category 'signalling'. Developmental processes include embryo and ovule development, seed germination, root and flower development while cellular responses to stimuli include response to heat, cold, metal ions, salt, osmotic and water stress. Signalling systems include heat and cold acclimation. All these processes are entirely consistent with the deduced activity of GmAC as an AC (Figure 1D), where cAMP acts as a second messenger (Frezza et al. 2018). When we extended the analysis to identify conditions that induce Glyma.07G251000 and its CEG50 (Table S1), we noted induction by various factors e.g., bud growth and salinity (Figure 2), consistent with its proposed function as an AC in early plant development and response to stress.

In order to practically investigate the inferred role of GmAC in stress response, Glyma.07G251000 was expressed in *E. coli* cells under different stress conditions, namely salt stress (200 mM NaCl) and oxidative stress (0.2 mM H_2O_2), followed by evaluation of cell growth by liquid and spot culture assays. In ordinary liquid media not having any stress condition, cells expressing the GmAC and those not expressing it, grew relatively the same, reaching a maximum growth density of around 1.25 nm (OD_{600}) after about 18 hours of incubation (Figure 3AI). However, the pattern of growth was noticeably different for cells growing in media with stress. In media with salt stress, cells expressing GmAC reached maximum growth after 20 hours of incubation with a culture density of around 1.0 nm compared to cells not expressing the protein that had a maximum growth density of 0.6 nm after a whole 24-hour period of incubation (Figure 3AII). 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 0.19 nm (Figure 3AIII). With regard to solid media, agar plates with no stress condition, similarly, showed no noticeable differences in colony sizes (~8.0 mm) between the *E. coli* cells expressing and not expressing GmAC and such an observation was made for all cellular dilution factors used in the experiment and after the complete incubation period of 24 hours (Figure 3BI). However, for agar plates with stress conditions, the picture was different. All colony sizes, at various dilution factors, of cells expressing the GmAC protein were noticeably larger (~6.0 mm) than those of cells not expressing the protein (~2.0 mm) whether grown under salt or oxidative stress (Figure 3BII-III). The ability of cells expressing the GmAC protein to grow under stress closely relates with the deduced inferred role of GmAC in stress response (Appendix A, Table S1) and induction by stress (Figure 2).

DISCUSSION

Soybean [*Glycine max* (L.) Merrill] is one of the most important oil crops of the world, which also has tremendous importance as a food legume. Soy oil finds a variety of uses for domestic and agricultural purposes in form of various food preparations and animal feeds (Probst and Judd 1973, Pratap et al. 2012). Besides its use for domestic and agricultural purposes, soy oil finds multifarious uses in industries and enterprises related to production of pharmaceuticals, plastics, papers, inks, paints, varnishes, pesticides and cosmetics (Song et al.

2011, Ko et al. 2013). Lately, use of soy oil as bio-diesel has opened up another possibility of renewable sources of energy for industrial uses. More so, as a leguminous crop, soybean is capable of utilizing atmospheric nitrogen through biological nitrogen fixation and is therefore, less dependent on synthetic nitrogenous fertilizers that have the risk of causing environmental pollution. However, despite all these benefits associated with this legume, cultivation of soybean is highly sensitive to various environmental challenges such as water shortage (Brevedan and Egli 2003), high temperatures (Liu et al. 2003) and microbial diseases (Deshmukh et al. 2014).

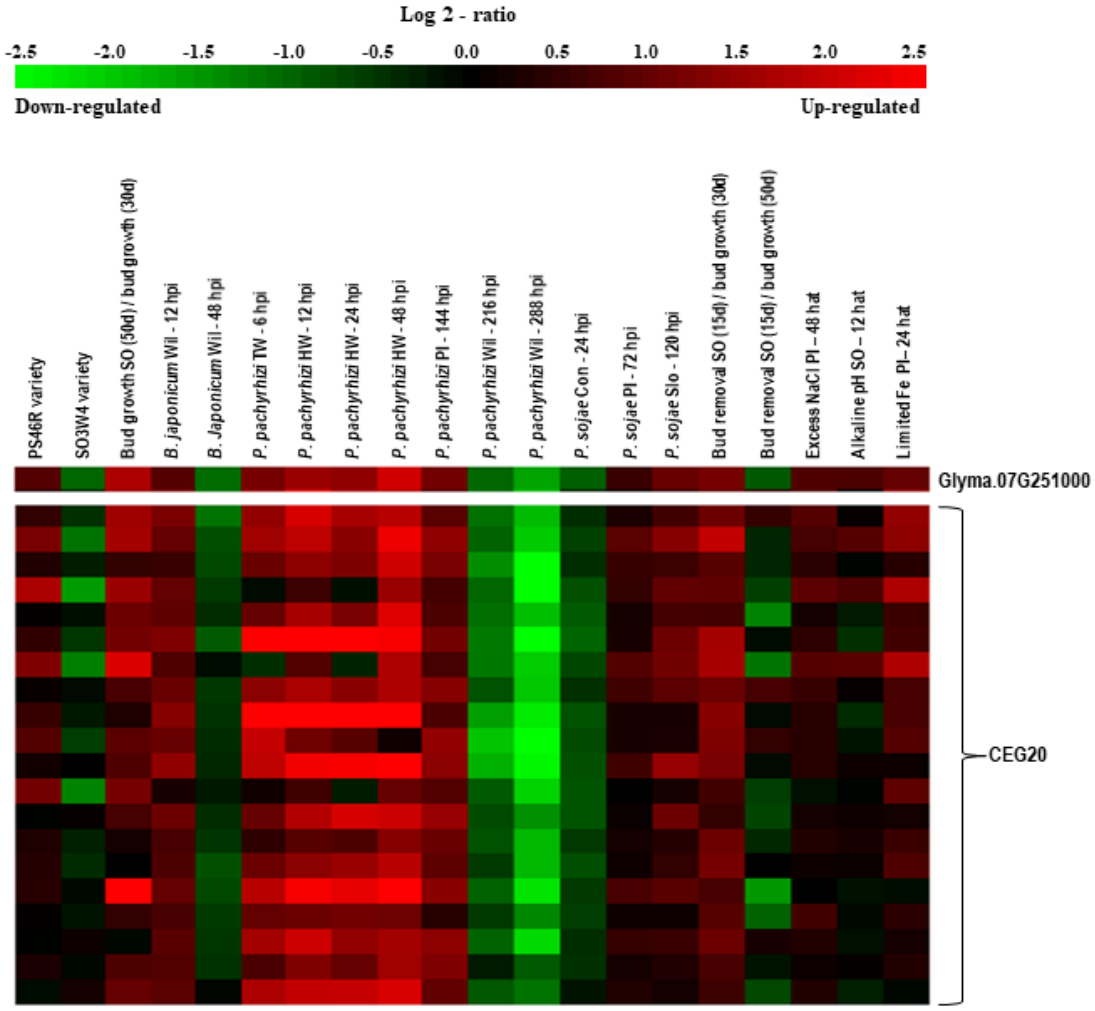

Figure 2. Heatmap constructed to illustrate the fold-change (log2) in expression of Glyma.07G251000 and 20 randomly selected expression-correlated genes (CEG20) in response to selected microarray experiments. The experiments presented include; untreated samples (PS46R and S03W4), bud growth (50d/30d, S03W4), *Botrytis japonicum* (12 and 48hpi, Williams), *Phakopsora pachyrhizi* (6, 12, 24, 48, 144, 216 and 288hpi, TW72-1, HW94-1, PI 459025B and Williams), *Phytophthora sojae* (24, 72 and 120hpi, Conrad, PI 291327 and SLO-1), bud removal/growth (15d/30d and 15d/50d, S03W4), salinity (48hat, PI 416937), alkalinity (12hat, S03W4) and limited iron (24hat, PI 416937), where d = days, hpi = hours post inoculation and hat = hours after treatment. Details of the microarray experimental conditions are presented in Appendix B.

Apparently, in plants and other organisms, response to stress or harsh conditions is mostly mediated by a special group of enzymes termed adenylyl cyclases (ACs) and their catalytic product, 3',5'-cyclic adenosine monophosphate (cAMP) (Steer 1976, Newton et al. 1999, Gehring 2010). However, while in a number of plant species, ACs have been identified and characterized (Moutinho et al. 2001, Ruzvidzo et al. 2013, Świeżawska et al. 2014, Al-Younis et al. 2015, 2018, 2021, Kasahara et al. 2016, Chatukuta et al. 2018, Bianchet et al. 2019, Ruzvidzo et al. 2019, Yang et al. 2021), in soybean, none has been identified yet. Thus any study targeted at understanding and elucidating the genome response of this legume to drought and other related environmental stress factors would be very important and worthwhile.

In *Arabidopsis thaliana*, the discovery of candidate ACs has largely been through a novel motif-based search, which involved identification of key amino acid residues in the catalytic center of known and experimentally tested nucleotide cyclases (NCs) (Ludidi and Gehring 2003, Gehring 2010). Using this systematic approach, various ACs

have been identified (Ruzvidzo et al. 2013, 2019, Chatukuta et al. 2018, Al-Younis et al. 2015, 2018, 2021, Bianchet et al. 2019, Yang et al. 2021), all harbouring the novel AC catalytic motif (Figure 1A) (Gehring 2010). Incidentally, in *Nicotiana benthamiana*, only one AC molecule (NbAC) has so far been identified (Ito et al. 2014) and this protein too, harbors the same annotated novel AC catalytic motif. In its identification, NbAC was phylogenetically clustered with various other protein molecules, which among them, included four of the previously confirmed ACs containing the novel AC search motif (Moutinho et al. 2001, Ruzvidzo et al. 2013, 2019, Chatukuta et al. 2018) and one uncharacterized soybean protein, GmAC, encoded by the Glyma.07G251000 gene (Ito et al. 2014). Guided by this nortion, we then targeted the GmAC protein in our study; firstly, to check if it harbors the AC catalytic motif; secondly, to establish if it is an AC, and lastly, to determine if it has a role in any key cellular process, e.g., stress response, of the soybean plant.

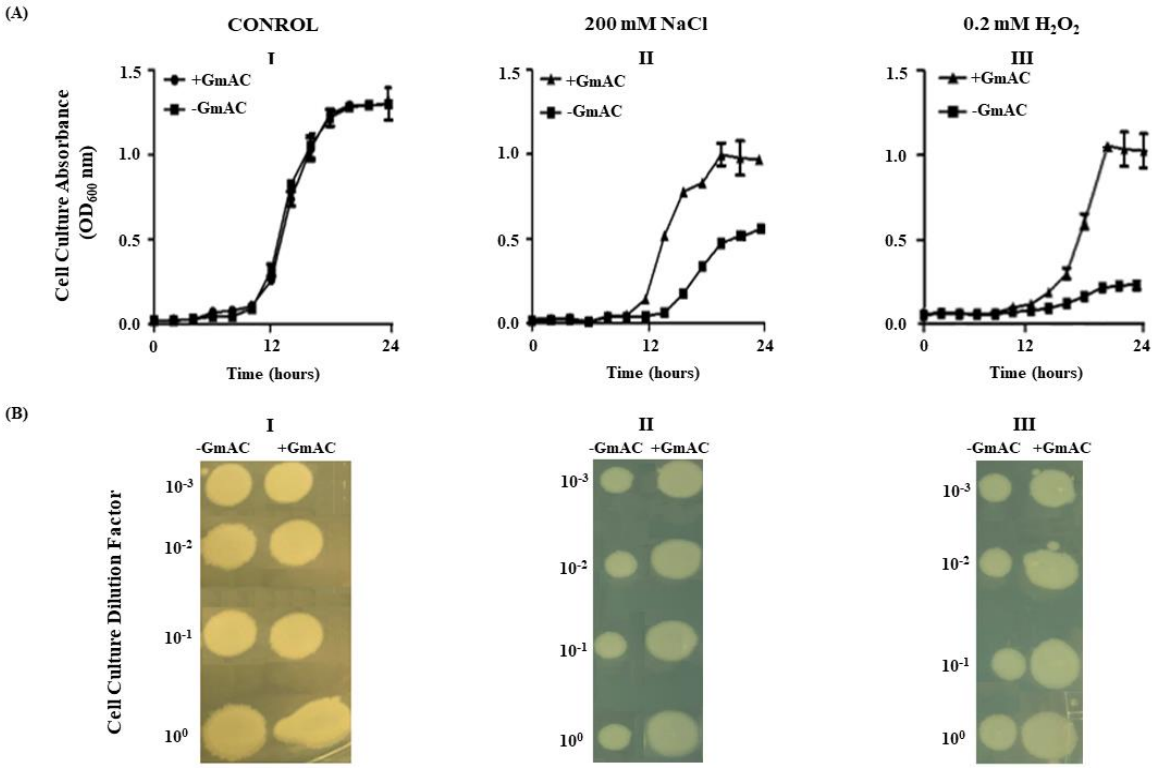

Figure 3. Determining the role of GmAC in stress response. *E. coli* cells expressing (+GmAC) and not expressing (-GmAC) the GmAC protein were cultured in (A) liquid and (B) solid media that was either ordinary (I) or supplemented with 200 mM NaCl for salt stress and (II), 0.2 mM H_2O_2 for oxidative stress (III), followed by evaluation of their growth patterns. For cells growing in liquid media, growth was measured through culture absorbances at OD_{600} 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. ($n = 3$ and $p < 0.05$ as determined by ANOVA).

When its amino acid sequence was retrieved followed by analysis by ExPASy (https://www.expasy.ch/), it was noted that, like the other several previously confirmed ACs (Moutinho et al. 2001, Ruzvidzo et al. 2013, 2019, Ito et al. 2014, Al-Younis et al. 2015, 2018, 2021, Chatukuta et al. 2018, Bianchet et al. 2019, Yang et al. 2021), GmAC contains the same novel AC catalytic motif (Figure 1B). We then cloned the Glyma.07G251000 gene and expressed its coded GmAC protein as a His-tagged fusion recombinant product of approximately 44.97 kDa (Figure 1C). When purified (Figure 1D, inset) and tested for AC activity *in vitro*, using enzyme immunoassay, recombinant GmAC showed a Mn^{2+}-dependent activity that is positively enhanced by Ca^{2+} (Figure 1D). Apparently, the ability of GmAC to exhibit a relatively higher (~5-fold) AC activity with Mn^{2+} than Mg^{2+} strongly points to this protein as a soluble AC (sAC) (Lomovatskaya et al. 2008) because sACs strictly prefer Mn^{2+} to Mg^{2+} as a co-factor of activity (Braun and Dods 1975) and are localized intracellularly (Kamenetsky et al. 2006). In *G. max*, GmAC is localized in the mitochondrion (http://www.genevestigator.ethz.ch) (Hruz et al. 2008) just like all other known sACs. Moreover, the activation of GmAC by calcium (~1.4 fold) then firmly confirms this protein as a sAC because only sACs, and not transmembrane ACs (tmACs), are functionally activated by the Ca^{2+} ion (Kamenetsky et al. 2006, Chen et al. 2000). Furthermore, the activation of GmAC by Ca^{2+} is consistent with the previously observed activation of plant ACs and GCs by the same metal ion (Steer and Levitzki 1975, Muleya et al. 2014, Wheeler et al. 2017, Chatukuta

et al. 2018, Ruzvidzo et al. 2019), whereby in GCs, Ca^{2+} acts as a switch between the kinase and GC activities of a dual-functioning phytosulfokine receptor (AtPSKR1) (Muleya et al. 2014). Apparently, the molecular mechanism through which this metal ion activates both the GCs and ACs is not currently known although it is probable that it invokes structural changes to the GC or AC catalytic center, thereby enhancing activity.

When expressed in *Escherichia coli*, GmAC complemented AC-deficiency (*cya*A mutation) in the mutant strain, SP850 (Figure 1E). This host strain lacks the only AC system available in *E. coli*, necessary for lactose fermentation (Ullmann and Danchin 1980, Shah and Peterkofsky 1991, Tang et al. 1995) and therefore, its rescue by any protein e.g., GmAC in this case, to ferment lactose, signifies the functional AC activity for such a protein. Besides GmAC, ten out of the eleven previously confirmed plant ACs have also demonstrated this capability of rescuing the SP850 host strain (Moutinho et al. 2001, Ruzvidzo et al. 2013, Świeżawska et al. 2014, Al-Younis et al. 2015, 2018, 2021, Kasahara et al. 2016, Chatukuta et al. 2018, Bianchet et al. 2019, Ruzvidzo et al. 2019, Yang et al. 2021).

In eukaryotes, since it is widely accepted that proteins that are co-expressed often have related functions and coordinated regulatory systems (Jansen et al. 2002, Allocco et al. 2004, Lee et al. 2004, Sweetlove and Fernie 2005), we then subjected the Glyma.07G251000 gene to an expressional correlation analysis so as to explore and gain insights into the probable biological functions of the GmAC protein in *G. max*. From this analysis, it emerged that the Glyma.07G251000 gene is mostly co-expressed with genes involved in embryo development and response to stress (Appendix A, Table S1), thus suggesting a core function for the GmAC protein in these two key cellular processes that are both mediated by cAMP (Gehring 2010). When the analysis was extended to identify conditions that induce expression of Glyma.07G251000 and its correlated genes (Appendix A, Table S1), we noted differential induction by bud growth (which represents the early plant development stage); *Botrytis japonicum* (a symbiotic bacterium necessary for nodulation during the juvenile stages of legume growth and nitrogen fixation thereafter) (Terakado et al. 1997, Colebatch et al. 2004, Kouchi et al. 2004, Brechenmacher et al. 2008); *Phakopsora pachyrhizi* and *Phytophthora sojae* (two major pathogens of soybean) (Ward 1990, Morris et al. 1998, Moy et al. 2004, Panthee et al. 2007); bud removal (which mimics herbivory) (Bi and Felton 1995, Gatehouse 2002, Kessler and Baldwin 2002); and salinity, alkalinity and nutrient deficiency (which are all abiotic stress factors) (Figure 2). All this was consistent with the proposed core function of GmAC in early plant development and response to stress.

To augment our findings, we used liquid and spot culture assays to demonstrate the role of Glyma.07G251000 and its gene product GmAC, in stress response. As is shown in Figure 3, GmAC conferred resistance to *E. coli* cells expressing it under salt stress (which mimics drought) and oxidative stress (which mimics pathogenesis). This again, corroborated fairly well with our results from the co-expression and stimulus-specific analyses of the Glyma.07G251000 gene (Table S1 and Figure 2) as well as the detection of AC activity in GmAC (Figure 1D).

In conclusion, our work has identified the first ever AC molecule in soybean with key role in early plant development and response to stress. Conceivably, findings of our study strongly give us extra interest to search for yet more ACs with similar or related functions to GmAC in *G. max*. This also lays an ideal platform for attempts to develop transgenic varieties of this legume that are AC-based (genetically-modified or over-expressers) with tolerance to stress. Such attempts would be very vital to the critical fields of legume breeding and production.

Acknowledgements

The project was funded by North-West University (NWU) and the National Research Foundation (NRF) of South Africa (Grant Numbers: CSUR78843 & CSUR93635).

References

Al-Younis I et al. 2015. The *Arabidopsis thaliana* K^+-uptake permease 7 (AtKUP7) contains a functional cytosolic adenylate cyclase catalytic centre. FEBS Letters 589:3848-3852.

Al-Younis I et al. 2018. The *Arabidopsis thaliana* K^+-uptake permease 5 (AtKUP5) contains a functional cytosolic adenylate cyclase essential for K^+ transport. Frontiers in Plant Science 9:1645.

Al-Younis I et al. 2021. Functional crypto-adenylate cyclases operate in complex plant proteins. Frontiers in Plant Science 12:1709.

Allocco D et al. 2004. Quantifying the relationship between co-expression, co-regulation and gene function. BMC Bioinformatics 5:18.

Bi JL and Felton GW. 1995. Foliar oxidative stress and insect herbivory: primary compounds, secondary metabolites, and reactive oxygen species as components of induced resistance. Journal of Chemical Ecology 21:1511-1530.

Bianchet C et al. 2019. An *Arabidopsis thaliana* leucine-rich repeat protein harbors an adenylyl cyclase catalytic center and affects responses to pathogens. Journal of Plant Physiology 232:12-22.

Braun T and Dods R. 1975. Development of a Mn^{2+}-sensitive, "soluble" adenylate cyclase in rat testis. Proceedings of the National Academy of Science, USA 72:1097-1101.

Brechenmacher L et al. 2008. Transcription profiling of soybean nodulation by *Bradyrhizobium japonicum*. The American Phytopathological Society 21:631-645.

Brevedan R and Egli D. 2003. Short periods of water stress during seed flling, leaf senescence, and yield of soybean. Crop Science 43:2083-2088.
Chai C et al. 2015. Soybean transcription factor ORFeome associated with drought resistance: a valuable resource to accelerate research on abiotic stress resistance. BMC Genomics 16:596.
Chatukuta P et al. 2018. An Arabidopsis clathrin assembly protein with a predicted role in plant defense can function as an adenylate cyclase. Biomolecules 8:E15.
Choi YE and Xu JR. 2010. The cAMP signaling pathway in *Fusarium verticillioides* is important for conidiation, plant infection, and stress responses but not fumonisin production. Molecular Plant-Microbe Interactions 23:522-533.
Colebatch G et al. 2004. Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. The Plant Journal 39:487-512.
Deshmukh H et al. 2014. The microbiota regulates neutrophil homeostasis and host resistance to *Escherichia coli* K1 sepsis in neonatal mice. Nature Medicine 20:524-530.
Frezza E et al. 2018. A molecular dynamics study of adenylyl cyclase: the impact of ATP and G-protein binding. PLoS One 13:e0196207.
Gatehouse J. 2002. Plant resistance towards insect herbivores: a dynamic interaction. The New Phytologist 156:145-169.
Gehring C. 2010. Adenyl cyclases and cAMP in plant signaling - past and present. Cell Commununication and Signalling 8:1-5.
Hruz T et al. 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. Advanced Bioinformatics 8:420747.
Ito M et al. 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:e27420.
Jansen R et al. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Molecular Microbiology 43:1565-1575.
Kamenetsky M et al. 2006. Molecular details of cAMP generation in mammalian cells: a tale of two systems. Journal of Molecular Biology 362:623-639.
Kasahara M et al. 2016. An adenylyl cyclase with a phosphodiesterase domain in basal plants with a motile sperm system. Scientific Reports 6:39232.
Kessler A and Baldwin I. 2002. Plant responses to insect herbivory: the emerging molecular analysis. Annual Reviews in Plant Biology 53:299-328.
Ko KP et al. 2013. Intake of soy products and other foods and gastric cancer risk: a prospective study. Journal of Epidemiology 23:337-343.
Kouchi H et al. 2004. Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. DNA Research 11:263-274.
Kwezi L et al. 2011. The phytosulfokine (PSK) receptor is capable of guanylate cyclase activity and enabling cyclic GMP-dependent signaling in plants. Journal of Biological Chemistry 286:22580-22588.
Lee HK et al. 2004. Co-expression analysis of human genes across many microarray data sets. Genome Research 14:1085-1094.
Liu F et al. 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 30:271-280.
Lomovatskaya L et al. 2008. Plant adenylate cyclases. Journal of Receptors and Signal Transduction Research 28:531-542.
Ludidi N and Gehring C. 2003. Identification of a novel protein with guanylyl cyclase activity in *Arabidopsis thaliana*. Journal of Biological Chemistry 278:6490-6494.
Minneci F et al. 2013. FFPred 2.0: improved homology-independent prediction of gene ontology terms for eukaryotic protein sequences. PLoS One 8:e63754.
Morris P et al. 1998. Chemotropic and contact responses of *Phytophthora sojae* hyphae to soybean isoflavonoids and artificial substrates. Plant Physiology 117:1171-1178.
Moutinho A et al. 2001. cAMP acts as a second messenger in pollen tube growth and reorientation. Proceedings of the National Academy of Sciences 98:10481-10486.
Moy P et al. 2004. Patterns of gene expression upon infection of soybean plants by *Phytophthora sojae*. Molecular Plant-Microbe Interactions 17:1051-1062.
Muleya V et al. 2014. Calcium is the switch in the moonlighting dual function of the ligand-activated receptor kinase phytosulfokine receptor 1. Cell Communication and Signaling 12:60.
Newton R et al. 1999. Cyclic nucleotides in higher plants: the enduring paradox. The New Phytologist 143:427-455.
Panthee D et al. 2007. Gene expression analysis in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) in an early growth stage. Functional and Integrated Genomics 7:291-301.
Pratap S et al. 2012. Can privatization improve animal healthcare delivery system? an ex-ante analysis of dairy farmers in tarai region of uttarakhand. Agricultural Economics Research Review 25:507-514.
Probst A and Judd R. 1973. Origin, U.S. history and development, and world distribution. In Soybeans: improvement, production, and uses. American Society of Agronomy 16.
Roelofs J et al. 2001. The Dictyostelium homologue of mammalian soluble adenylyl cyclase encodes a guanylyl cyclase. EMBO Journal 20:4341-4348.
Ruzvidzo O et al. 2013. Recombinant expression and functional testing of candidate adenylate cyclase domains. In Cyclic nucleotide signaling in plants: Methods and Protocols. vol 1016. pp. 13-25.
Ruzvidzo et al. 2019. New perspectives on plant adenylyl cyclases. Frontiers in Molecular Biosciences 6:136.
Saeed A et al. 2003. TM4: a free, open-source system for microarray data management and analysis. BioTechniques 34:374-378.
Schmutz J et al. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463:178-183.
Shah S and Peterkofsky A. 1991. Characterization and generation of *Escherichia coli* adenylate cyclase deletion mutants. Journal of Bacteriology 173:3238-3242.
Song F et al. 2011. Biodegradable soy protein isolate-based materials: a review. Biomacromolecules 12:3369-3380.
Steer M. 1976. Cyclic AMP. Annals of Surgery 184:107-115.
Steer ML and Levitzki A. 1975. The control of adenylate cyclase by calcium in turkey erythrocyte ghosts. Journal of Biological Chemistry 250:2080-2084.
Sweetlove L and Fernie A. 2005. Regulation of metabolic networks: understanding metabolic complexity in the systems biology era. New

Phytology 168:9-24.
Świeżawska B et al. 2014. Molecular cloning and characterization of a novel adenylyl cyclase gene, HpAC1, involved in stress signaling in *Hippeastrum hybridum*. Plant Physiology and Biochemistry 80:41-52.
Tang W et al. 1995. Truncation and alanine-scanning mutants of type I adenylyl cyclase. Biochemistry 34:14563-14572.
Terakado J et al. 1997. Cyclic AMP in rhizobia and symbiotic nodules. Annals of Botany 80:499-450.
Tucker C et al. 1998. Two amino acid substitutions convert a guanylyl cyclase, RetGC-1, into an adenylyl cyclase. Proceedings of the National Academy of Science, USA 95:5993-5997.
Turner M et al. 2012. Genome organization and characteristics of soybean microRNAs. BMC Genomics 13:169.
Ullmann A and Danchin A. 1980. Role of cyclic AMP in regulatory mechanisms in bacteria. Trends in Biological Sciences 5:95-96.
Ward E. 1990. The interaction of soya beans with *Phytophthora megasperma* f.sp. glycinea: pathogenicity. Paper presented at the Biological Control of Soil-borne Plant Pathogens, Kyoto, Japan.
Wheeler J et al. 2017. The brassinosteroid receptor BRI1 can generate cGMP enabling cGMP-dependent downstream signaling. Plant Journal 91:590-600.
Yang H et al. 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 72:283-301.
Yunghans W and Moore D. 1977. Adenylate cyclase activity not found in soybean hypocotyl and onion meristem. FAO (AGRIS) 60: 144-149.