Molecular characterization and differential expression of an aromatic heptaketide producing type III plant polyketide synthase from Himalayan rhubarb

Shahzad A. Pandith1,2 · Niha Dhar2,3 · Sumedha Bhosale4 · Vitthal T. Barvkar5 · Sumeer Razdan2,6 · Manzoor A. Shah1 · Surrinder K. Lattoo2

Received: 27 July 2021 / Revised: 14 November 2021 / Accepted: 2 January 2022 / Published online: 15 January 2022
© Korean Society for Plant Biotechnology 2022

Abstract
Rheum australe (Himalayan Rhubarb, Polygonaceae), an endangered medicinal and vegetable herb owes its age-old rem-edying properties to the bio-active phyto-constituents viz. anthraquinones, stilbenoids, chromones and dietary flavonoids. Polyketide pathway primarily involving the intricate Type III polyketide synthases (PKSs) contributes to the biosynthesis of these phyto-constituents. In the present study, we perform a homology-based approach to isolate an 1176 bp full-length cds sequence of the RaALS gene showing an equitable level of sequence similarity to related Type III PKSs at both nucleic acid and amino acid levels. In silico characterization revealed the presence of highly conserved amino acid residues found in nearly all Type III PKSs including the conserved active-site residues, signature motif and cyclization pocket residues with an exception of Ile256 and Gly258. Docking studies established major interactions between the starter acetyl-CoA and RaALS. Copy number analysis suggested slender evolution in Type III PKS in R. australe having a single copy of RaALS gene. qRT-PCR analyses revealed corroboration between the higher expression of RaALS in leaves followed by stem and root with that of the metabolite concentration. Expression studies further showed a direct increase of RaALS transcripts with the growing metabolite accretion in relation to altitude suggesting a probable involvement of specific Type III PKS in biosynthesis of the major phyto-constituents. Furthermore, abiotic stressors viz. methyl jasmonate, salicylic acid and UV light enhanced RaALS transcription hinting towards its role in defense mechanism in R. australe and highlighting the sig-nificance of RaALS as a prospective target for metabolic engineering.

Keywords Aloesone synthase · Heterologous expression · Molecular docking · Promoter analysis · qRT-PCR expression · Rheum australe · Southern blotting
Introduction

Polyketide natural products (PNPs), a large family of chemical constituents are known for numerous biological as well as physiological functions in microbes and plants (Funabashi et al. 2008; Dao et al. 2011; Zeng et al. 2012) in addition to various pharmacological activities extensively reviewed previously (Pandith et al. 2018, 2020). The implication of these metabolites is copiously demonstrated by varied compounds viz. erythromycin, lovastatin, rapamycin, streptomycin, besides the low molecular weight anthraquinones, stilbenoids, and flavonoids (Flores-Sanchez and Verpoorte 2009; Nair et al. 2012; Pandith et al. 2020). In fact, various plants producing these PNPs have garnered extensive research interest in terms of their molecular aspects, biosynthetic pathways, structure and mode of action as they exhibit both structural and functional disparity.

Himalayan Rhubarb (Rheum australe, family Polygonaceae) is one such high-value medicinal herb of Northwest Himalayas which holds wide pharmacological significance due to the presence of bio-active PNPs as evidenced by various traditional medical systems viz. Chinese, Unani, Ayurvedic, etc. In fact, the swelling demand, nearly negligible cultivation and reckless harvesting from the wild has rendered the plant threatened (Pandith et al. 2014). The therapeutic properties of this broadly used high-altitude perennial herb are ascribed to important class of secondary chemical constituents (PNPs) which mostly include anthraquinones, chromones and stilbenoids, besides the potent phytotoxic flavonoids (Jiang et al. 2015; Pandith et al. 2020).

PNPs in plants are reported to be biosynthesized through successive condensation of simple malonyl-CoA-derived two-carbon acetate units with an acyl starter (CoA thioester) by multifunctional enzymes known as polyketide synthases (PKSs) categorized into Type I, II and III based on their architectural configurations. The Type III PKSs are iteratively acting and condensing homodimeric (80–90 kDa in native homodimer state) enzymes with a single KS (ketosynthase) domain which performs all the functions as executed by the important domains of Type I and Type II PKSs (Hertweck 2009). Studies have shown that the Type III PKSs exhibit their existence in nearly all the studied forms of life, barring animal (metazoan) and archaeal genomes, in a lineage-specific manner with at least a single representative of them that may possibly have paralogous forms as well. Pertinently, aloesone synthase (RaALS) is one of the important members of Type III PKS gene family which is known to catalyze six polyketide extensions of the starter acetyl-CoA. The enzyme is known to share about 60% identity with orthologous PKSs while maintaining the conserved catalytic triad (Cys164, His303 and Asn336) and Phe216—numbering in Alfalfa CHS. Nonetheless, unlike the well-characterized CHS and CHS-like enzymes, RaALS produces a heptaketide intermediate that undergoes decarboxylative C8–C13 aldol cyclization with an additional second ring formation to ultimately result in the formation of aloesone (2-acetyl-7-hydroxy-5-methylchromone), a biosynthetic precursor of the anti-inflammatory agent aloesin (aloesone 8-C-β-D-glucopyranoside) (Fig. 1). Interestingly, one of the distinctive features of RaALS which somehow separates it from the CHS family members is the respective replacement of three conserved residues (Thr197, Gly256, and Ser338; numbering in Medicago sativa CHS) with Ala, Leu, and Thr which affect the shape/volume of the initiation/elongation cavity of the active site. However, this ~42 kDa protein is generally known to maintain the overall architecture (fold) as that of the CHS superfamily enzymes (Abe et al. 2006; Lim et al. 2016).

As the major bio-active compounds are synthesized by Type III PKSs like CHS and CHS-like enzymes, the role of RaALS (a member of CHS-like enzyme group) has been anticipated in influencing the biosynthesis of these metabolites in R. australe. It therefore makes it imperative to investigate the positioning, expression profile and regulatory framework of RaALS gene to decipher its potential as the eventual target for metabolic engineering. Later would also aid in homologous/heterologous intensification of pharmacologically and commercially significant PNPs in R. australe. In this context, we characterized this heptaketide producing Type III PKS from R. australe for the first time as no significant reports of its effects on the overall growth and development of the plant are reported. It is rather essential to determine the role of RaALS in various biochemical and eco-physiological aspects of the plant. Indeed, we have previously reported the functional validation of two CHS paralogs exhibiting promiscuous behavior with varied CoA thioester substrates to yield the final product naringenin-chalcone/naringenin (Pandith et al. 2016).

In the present study, we describe the cloning and characterization of RaALS, a critical enzyme for the biosynthesis of chromones, from R. australe. The full-length gene was cloned and further expressed in the alternate/heterologous microbial (bacteria and yeast) hosts. The multiple sequence alignment and phylogenetic analyses were performed to determine the evolutionary relatedness of deduced RaALS sequence with the related Type III PKSs. Additionally, homology modeling and molecular docking studies were executed to confirm and validate the protein structure, and to gain insights into the donor and acceptor interactions at molecular level. The spatial (different tissues) and altitudinal (different locations along an altitudinal gradient) expression profiling of the transcripts
of RaALS was also studied. Moreover, the exogenous elicitors viz. methyl jasmonate (MeJA), salicylic acid (SA) and UV light, selected on the basis of isolated cis-acting promoter elements, were also evaluated in in vitro established cultures to study the expression profile of RaALS gene. Further, the gene copy number of the enzyme was determined using Southern blot technique. So far, this is the only report of RaALS characterization from the medicinal herb R. australe implicating its possible role in the regulation and biosynthesis of species-specific bio-active phyto-constituents. This work, therefore, emphasizes the molecular features of RaALS that can be potently exploited using various metabolic engineering and system biology approaches to enhance the survival of R. australe, an endangered herb against various stresses using cues from elicitation data. Additionally, highlighting RaALS, as a prospective target for enhancement of secondary metabolite production in planta for commercial purposes would be imperative.
Methodology

Plant material

The source materials were the cultures of *R. australe* those were established in vitro. The cultures were established from germinated saplings maintained under greenhouse conditions at Indian Institute of Integrative Medicine, CSIR, Jammu, India (32° 44′ N longitude, 74° 55′ E latitude; 305 m asl) as mentioned in the previous study (Pandith et al. 2016). In fact, a well-developed and reproducible in vitro regeneration system of *R. australe* with varying concentrations of several phyto-hormones was developed (Table S1). These in vitro raised cultures were maintained in Murashige and Skoog medium and further used for the elicitor (MeJA, SA) and UV-B treatments as external stressors to study their effect on the expression of *RaALS* gene (Fig. 2).

RNA extraction and cDNA synthesis

Fresh or cryopreserved samples were used for total RNA isolation as described earlier (Pandith et al. 2016). The isolated RNA was incubated for 30 min at 37 °C with the enzyme DNase I (Fermentas, Burlington, Canada) which removes possible traces of genomic DNA from the sample. Further, the quality of RNA extracted was assessed by running the sample on 1% formaldehyde agarose gel, and also by determining the absorbance ratio (A₂₆₀/₂₈₀) using spectrophotometer (AstraAuriga, Cambridge, UK).

For cDNA synthesis, 3–5 μg of DNase I treated total RNA was reverse-transcribed using Revert-aid premium reverse transcription kit (Fermentas, Burlington, Canada) with a modified Adapter-oligo-(dT) primer according to the manufacturer’s protocol. Briefly, 3–5 μg of purified RNA was mixed with 10 μM oligo(dT) primer and subjected to denaturation at 65 °C for 5 min followed by chilling on ice for 2 min. The reverse transcription reaction was carried out in a final volume of 20 μl reaction which contained 1× first-strand buffer (250 mM Tris–HCl, pH 8.3; 250 mM KCl; 20 mM MgCl₂; 50 mM DTT), 10 mM dNTPs and 1 μl of Moloney murine leukemia virus reverse transcriptase (200 units/μl). The reaction was incubated at 42 °C for 60 min followed by 70 °C for 5 min to inactivate the enzyme. The cDNA was stored at –20 °C for further use.

Cloning of *RaALS*

Degenerate primers (Table 1) were designed following the use of Blastn/Blastx (Sayers et al. 2021) and ClustalOmega (Sievers and Higgins 2021) programmes. The primer sequences were based on highly conserved regions of nucleotide sequences of Type III plant polyketide synthases (PKSs) retrieved from the GenBank™ data base at NCBI (National Centre for Biotechnology Information). Using synthesized cDNA as template, RT-PCR (reverse transcriptase polymerase chain reaction) for core amplification of *RaALS* was carried out with the following optimized cyclic conditions: 1 cycle of 94 °C for 3 min; followed by 35 cycles of 94 °C for 30 s, 60 °C for 45 s and 72 °C for 1:30 min; and a final extension of 72 °C for 10 min in a thermal cycler (Bio-Rad Laboratories, Hercules, CA, USA). The screened amplicons of *RaALS* were cloned into pTZ57R/T vector (Fermentas, Burlington, Canada), transformed into an *E.
coli host strain (DH5™; Invitrogen, Merelbeke, Belgium) and sequenced (ABI PRISM® 3130XL genetic analyzer; Applied Biosystems, Foster City, CA, USA). Sequences obtained were analyzed using the similarity search Blastn (Sayers et al. 2021) programme to ensure homology and were further used to design gene-specific primers (GSPs) for RACE PCR.

5’ and 3’ RACE

5’ and 3’ RACE was performed using Gene Racer cDNA amplification kit according to the instructions given in product manual (Invitrogen, USA). Respective cDNAs obtained were used to generate the flanking regions of the core amplicons from either side separately in two sets of PCRs. The primary reaction was performed using 5’/3’ RACE adapter primer (5’/3’ RACE_OUT) and 5’/3’ ALS_OUT as GSPs, whereas the secondary reaction was carried out using inner adapter primers (5’/3’ RACE_INN) and 5’/3’ ALS_INN as GSPs while using amplified products from first reaction set as template. All the PCR reactions were carried out in a 50 μl reaction volume containing 1 μl cDNA as template (except for secondary reaction which uses amplified products of initial reaction as template), 2.5 μl each of 10 μM adapter primers and GSPs for respective reactions and 44 μl of master mix (33.5 μl PCR grade water; 10 mM Tris–HCl, pH 9.0; 50 mM KCl; 2.5 mM MgCl2; 0.2 μM dNTPs and 2.5 U of Taq DNA polymerase). The reaction thermo-profile for both initial and nested PCR amplifications was as follows: 3 min at 94 °C, 35 cycles (30 s at 94 °C, 30 s at 60–65 °C, 1:30 min at 72 °C) and 10 min at 72 °C followed by hold step at 4–15 °C. The obtained 5’ and 3’ nested amplicons were purified and sub-cloned into pTZ57R/T vector and further sequenced. The sequences acquired from core, 5’ and 3’ RACE fragments were refined (trimmed to eliminate bad quality bases with low confidence scores, QV < 20), aligned and subsequently analyzed using Blastn/BlastX (Sayers et al. 2021) tools to validate the prediction of targeted ALS.

Full-length cloning of RaALS

The sequences of core fragments and 5’/3’ RACE products were compared and aligned to generate the full-length cDNA of RaALS which was then amplified using full-length primers viz FulALS_F and FulALS_R (Table 1). A high-fidelity proof-reading DNA polymerase (New England Biolabs, Herts, UK) was used for amplification of complete ORF of RaALS while using the respective primers. The reaction thermo-profile was as follows: 1 cycle for 3 min at 94 °C; followed by 35 cycles of 94 °C for 30 s, 60 °C for 45 s, and 72 °C for 1:30 min; and a final extension at 72 °C for 10 min.

| Cis-element | Position (strand) | Signal sequence | Putative function |
|-------------|------------------|-----------------|-------------------|
| ARE         | 33 (−)           | AAACCA          | Cis-acting regulatory element essential for the anaerobic induction |
| Box I       | 58 (+)           | TTTCAA          | Light-responsive element |
| CAAT box    | 153 (+), 220 (+), 261 (−), 262 (−), 380 (−), 381 (−) | CAAT, CAATT     | Common cis-acting element in promoter and enhancer regions |
| CCGTCC box  | 125 (+)          | CCGTCC          | Cis-acting regulatory element related to meristem specific activation |
| CGTCA motif | 217 (+)          | CGTCA           | Cis-acting regulatory element involved in the MeJA-responsiveness |
| GATA box    | 2 (+), 186 (−), 190 (−), 320 (−), 340 (−), 394 (−) | GATA            | Required for high level, light regulated, and tissue-specific expression |
| TATA box    | 254 (+), 256 (+), 313 (−), 312 (+), 314 (+), 255 (−), 258 (−) | TATA, TATAA, ATT, TATAA, TTTTA | Core promoter element around -30 of transcription start |
| TC-rich repeats | 71 (+)     | ATTCTCTAAC      | Cis-acting element involved in defense and stress responsiveness |
| TGACG motif | 217 (−)          | TGACG           | Cis-acting element involved in transcriptional activation of several genes by auxin and/or salicylic acid |
| MYB core    | 25 (+), 77 (−), 95 (−), 99 (+), 358 (−) | CNGTTR          | Binding site for all animal MYB and at least two plant MYB proteins; ATMYB1 and ATMYB2, both isolated from Arabidopsis |
| Skn-1_motif | 83 (+), 119 (−)   | GTCAT           | Cis-acting regulatory element required for endosperm expression |
| GT1 motif   | 357 (−)          | GGTGAA          | Light-responsive element |

Table 1: Putative cis-acting regulatory elements identified in the promoter of RaALS using PLANTCARE database.
Full-length ORF was ligated into blunt-ended pJET vector and consequently sub-cloned into E. coli DH5α.

Computational and phylogenetic analyses of RaALS sequence

The complete cds fragment obtained was translated using ExPASy Translate tool (http://web.expasy.org/translate/), secondary structures predicted by SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) and the properties of deduced amino acid sequence were estimated using ExPASy ProtParam (http://web.expasy.org/protparam/) and Compute pi/Mw (https://web.expasy.org/compute_pi/) online tools. SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP-4.1/) was used to determine the presence of signal peptides and transmembrane helices predicted by TMHMM v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) server. The amino acid sequences were aligned using Clustal Omega Multiple Sequence Alignment tool (Sievers and Higgins 2021) and the identification of functionally and structurally important residues in protein sequences was done using ConSeq and ConSurf programmes (Ashkenazy et al. 2010). The deduced amino acid sequence of RaALS was aligned with multiple CHS plant sequences using Clustal W embedded in MEGA X (Kumar et al. 2018). Maximum Likelihood Method was used to construct the phylogenetic tree with MEGA X software with 500 bootstrap replicates to obtain confidence levels with the branches.

Homology modeling and model validation

The primary sequence of RaALS was retrieved from UniProt (Accession number A0A0A0P5P9). Crystal structure of CHS1 (PDB ID: 4YJY, Resolution 1.86 Å) was chosen as template with 64% sequence identity and 99% query coverage. Homology model was built using the SWISS-MODEL online server (Waterhouse et al. 2018). Model was energy minimized using Chimera version 1.13 (Pettersen et al. 2004) and amber ff12SB force field was used for parameterization. Indeed, hundred runs of steepest descent followed by conjugate gradient cycles were performed for energy minimization and QMEAN score was computed to analyse the overall quality of models (Benkert et al. 2011). The generated model was further validated by different structure assessment tools. Ramachandran plot was generated using PROCHECK tool (Laskowski et al. 1993) while ProSA (Wiederstein and Sippl 2007) made plot of local model quality. Further, molecular graphics and RMSD analyses were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

Docking simulations of substrate and product molecule to aloesone synthase

Model structure was subjected to protein–ligand docking by AutoDock Tools version 4.2 (Morris et al. 2009). Starter molecule acetyl-CoA and product aloesone were docked to the protein model. Prior to molecular docking, all hydrogen atoms were added to the protein structure followed by Gastegier charges. AutoGrid 4.2 module of AutoDock 4.2 was used to produce grid maps. Hundred independent runs with step size of 0.2 Å for translations and 5° per step for torsions and initial population of random individuals with population sizes of 300 individuals were set as fixed parameters for all the docking analyses using genetic algorithm (GA). For GA, the maximum number of energy evaluation was set to 2,500,000, maximum number of generations was set to 27,000, and maximum number of top individual that automatically survived was set to 1 with the mutation rate of 0.02 and crossover rate of 0.8. The Lamarckian Genetic Algorithm was chosen for generating the best conformer and other docking parameters were set as default. Docked complexes were visualized by PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.

Plasmid construction and heterologous expression in E. coli and yeast

Full-length coding sequence of RaALS gene was modified by adding restriction sites. Both forward (FulALS_F) and reverse (FulALS_R) primers were engineered to introduce BamHI and EcoRI restriction sites at the beginning and end of the coding sequence, respectively. The primers were tailored for directional cloning in pGEX4T-2 vector and heterologous expression in E. coli. ORFs containing modified restriction sites were cloned and excised from pJET vector (Fermentas, Vilnius, Lithuania) with BamHI and EcoRI and reconfirmed by sequencing before sub-cloning into the respective restriction sites of pre-digested and purified bacterial expression vector pGEX-4 T-2. The cloned RaALS was expressed as fusion protein with GST tag at N-terminus of the hybrid expression vector. Further, the heterologous expression of the recombinant protein was carried out as described by us earlier (Pandith et al. 2016). Briefly, the pGEX-RaALS expression cassette was transformed into E. coli BL21 (DE3) cells (Invitrogen, Merelbeke, Belgium) and single colony from the recombinant culture was cultured in Luria–Bertani (LB) medium containing ampicillin (100 mg/ml) at 37 °C on a shaker at 200 rpm for overnight. 1% culture was inoculated in fresh LB medium with respective antibiotic and incubated at 37 °C until optical density (A600) reached 0.4–0.6. The protein expression was induced by adding varied concentrations (0.2–1 mM) of isopropyl β-D-1-thiogalactopyranoside (IPTG; Fermentas, Burlington, Belgium)
Canada) into the cultures. The cells growing at same condition without IPTG induction were used as control. The cultures were constantly incubated at 30 °C for 8–12 h and cells were harvested after every two hours by centrifugation. The cell pellets obtained were re-suspended in 6X SDS loading buffer (0.375 M Tris, pH 6.8; 12% SDS; 60% glycerol; 0.6 M DTT; 0.06% bromophenol blue) and the expression was determined by running the samples on 10% sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE).

Additionally, RaALS gene was also expressed in the fission yeast, Schizosaccharomyces pombe as described in one of the earlier investigations from our lab (Dhar et al. 2014). BglII restriction endonuclease (Table 1) was employed to design, clone (directional cloning), digest and further express the RaALS gene in S. pombe using the expression vector pDS472a under the control of nmt1 promoter.

### Genomic southern blot analysis

The genomic DNA was isolated from mature leaves of R. australis using the standard cetyl-trimethyl-ammonium bromide (CTAB) method with slight modifications [4% PVP was used in the buffer instead of 1%, and one P:C:I (25:24:1) and two C:I (24:1) treatments were given to the supernatant remaining from lysed tissue] to increase the yield of DNA extracted (CTAB DNA extraction buffer 2009). About 20 µg of extracted, purified and quantified (at A260/280 nm) DNA was digested for 16–20 h at 37 °C with restriction endonucleases BamHI, EcoRI (non-cutters) and NolI (single cutter). The digested product was separated by electrophoresis on 0.7–0.8% (w/v) agarose gel in Tris–borate/EDTA buffer and then transferred onto a positively charged nylon membrane (Roche, Germany) and hybridized with the digoxigenin (DIG)-labeled DNA probe as described in our previous studies (Rana et al. 2013; Dhar et al. 2014; Pandith et al. 2016). The probe for RaALS was synthesised by PCR using full-length primers (FulALS_F and FulALS_R; Table 1) that amplified the coding sequence of the gene. The steps of probe labelling, hybridization, blocking, washing, and signal detection were done as per the instructions given in the user manual of the DIG-DNA labelling and detection kit (Roche Diagnostics GmbH, Mannheim, Germany).

### Tissue-specific and site-specific transcript study

The tissue- and site-specific expression profiling of RaALS was determined by qRT-PCR (quantitative real-time PCR) analysis. To study this relative expression, total RNA was isolated from different plant parts viz. root, stem and leaf, and from the leaf tissues of plant samples collected from four different geographic locations (L1, L2, L3 and L4) of northwestern Himalayas as reported earlier (Pandith et al. 2014). A total of three replicates were taken for each sample. For every sample, 3 µg of DNase-treated RNA was reverse-transcribed using the iScript cDNA synthesis kit (Bio-Rad) to synthesize the first-strand cDNA as per the product manual. The qRT-PCR reactions were performed in triplicates following SYBR-based chemistry (Takara, Japan) while using the ABI StepOne real-time quantitative PCR system (Applied Biosystems, USA). Concisely, the standard reaction contained 0.5 µl of cDNA as template, 200 µM each primer (Table 1), 5 µl of SYBR PremixExTaq and nuclelease-free water to make up the final volume to 10 µl. The manufacturers’ manual was followed for cycling conditions: holding stage of one cycle at 95 °C for 10 min, followed by the cycling stage (40 cycles) of 95 °C for 15 s and 60 °C for 1 min, and finally melting curve stage of 95 °C for 15 s, 60 °C for 1 min, and 95 °C for 15 s. The real-time primers were designed by Primer Express version 3.0 (Applied Biosystems, Foster City, CA) and were further confirmed by a dissociation curve (observation of a single peak for each primer pair). A housekeeping gene β-actin, amplified with Actin_F and Actin_R primers (Table 1), was used as an endogenous control to normalize the expression of RaALS. The qRT-PCR data generated were analyzed on the basis of comparative Ct (cycle threshold) method as described earlier (Pandith et al. 2016).

### Promoter isolation

Genome walking approach was followed to determine the promoter sequence of RaALS gene using GenomeWalker Universal Kit (Clontech, Palo Alto, CA, USA). Genomic DNA was isolated from fresh and young leaves of R. australis using DNeasy Plant mini Kit as per the instructions of the manufacturer (Qiagen, Hilden, Germany). To construct GenomeWalker DNA libraries, the genomic DNA was further digested with four different blunt-end-generating restriction endonucleases (DraI, PvuII, EcoRV, and StuI) in separate tubes each containing 5 µg of the extracted DNA. The digested samples were recovered using a PCR purification kit (Qiagen) and ligated individually to the GenomeWalker AP adaptor (provided with kit) to produce four adapter-ligated libraries to be used as templates in PCRs consisting of two walking amplification steps (primary and nested PCR). Two-round PCR reactions were carried out using AP1 and AP2 adapter primers (provided with kit) and the gene-specific primers (Table 1) in light of the kit manual with slight modifications. The final products generated from this exercise were analyzed on agarose gel (1.5–2%), purified and further cloned into pTZ57R/T cloning vector for sequencing to obtain the promoter sequence of RaALS gene. The sequence thus generated was processed and then subjected to the PlantCare (Lescot et al. 2002) and New PLACE (Higo et al. 1999) databases to identify various
putative cis-acting elements in the promoter region of this specific Type III PKS.

Elicitor treatment and qRT-PCR expression analyses of \textit{RaALS}

In vitro cultures of \textit{R. australe} were developed to determine the variation in the accumulation pattern of \textit{RaALS} transcripts upon treatments with different plant-derived endogenous elicitors as discussed in our previous study (Pandith et al. 2016). Briefly, the axenic cultures were adapted for 2 weeks in MS (Murashige and Skoog) liquid medium before subjecting to elicitations by MeJ (methyl jasmonate, 0.1 mM) and SA (salicylic acid, 0.1 mM) and UV-B radiation. The elicitors were chosen in correspondence to the promoter motifs identified in the upstream region of the \textit{RaALS} gene. The control and treated plant tissue samples were harvested for RNA extraction at defined time periods of 0 (in case of control), 12, 24 and 48 h, except for UV-B exposed samples which were harvested at 3, 6 and 9 h. The RNA samples from control and treated samples were then reverse-transcribed to synthesise cDNA and the effects of elicitor treatments on the expression analysis of \textit{RaALS} gene were studied using qRT-PCR with the same parameters as used for the tissue- and site-specific transcript study discussed above.

**Statistical analyses**

Data were analysed using GraphPad Prism for Windows Version 5.0 (GraphPad Software, La Jolla California USA; \url{www.graphpad.com}) by one-way ANOVA followed by Tukey’s test \((P < 0.05)\) to determine the differences among means.

**Results and discussion**

**Molecular cloning of \textit{RaALS}**

A homology-based approach was used to isolate and identify the ALS gene from \textit{R. australe} while taking benefit from some of the highly conserved amino acid sequences of Type III plant polyketide synthases. First, a 319 bp core cDNA fragment was obtained by PCR using degenerate primers (Table S2) which was subsequently followed by 5’ and 3’ RACE generating 496 and 361 bp fragments, respectively. Assemblage of all these fragments resulted in the generation of an 1176 bp full-length cds sequence of the \textit{RaALS} gene. The obtained sequence (ORF) was submitted to the NCBI GenBank database under the accession number KC473812. The nucleic acid sequence alignment showed a reasonable level of sequence similarity to related Type III PKSs using BLASTN/BLASTX analysis tools. Additionally, using the BLASTX/BLASTP algorithm, the \textit{RaALS} amino acid sequence displayed extended sequence similarity (64–88%) with orthologous sequences from other plant species belonging to the order Caryophyllales. To cite a few: \textit{Chenopodium quinoa} (GenBank accession no. ATY35195.1), \textit{Drosophyllum lusitanicum} (GenBank accession no. ABQ59603.1), \textit{Fagopyrum tataricum} (GenBank accession no. AHA14502.1), \textit{Fallopia multiflora} (GenBank accession no. ATY35195.1), \textit{Polygonum cuspidatum} (GenBank accession no. ABE68892.1), and \textit{Spinacia oleracea} (GenBank accession no. XP_021835131.1). This kind of similarity index showed that \textit{RaALS} gene is a prime member of the Type III PKS family, and is supposed to behave in a similar fashion in both structural and functional aspects. Moreover, as a prerequisite for heterologous expression, and to further attempt its functional validation in an alternate host, an orthologous CHS-like gene was successfully isolated from \textit{R. australe}.

**In silico characterization and phylogenetic analysis of deduced \textit{RaALS}**

The ORF of \textit{RaALS} was subjected to translation to generate 391 amino acids, corresponding to protein of 43.35 kDa with calculated pI value of 5.74. The secondary structure analysis revealed that \textit{RaALS} are predominantly α-helical protein with respective percentages for α-helices (43.99%, 172 residues), random coils (34.27%, 134), extended strands (15.09%, 59 residues) and β-turns (6.65%, 26 residues). The computed instability index (II) was found to be 37.26 confirming the stability of the protein secondary structure. Further, the aliphatic index of 83.81 and grand average of hydropathicity (GRAVY) of −0.157 indicates toward the hydrophobic nature of the protein. Pertinently, and as confirmed by SignalP 4.1 and TMHMM v. 2.0 servers, \textit{RaALS} was also found to lack any signal peptides and, thus, the transmembrane helices normally associated with the hydrophobic proteins. Further, analysis of the evolutionary conservation of \textit{RaALS} amino acid sequence revealed various high-score structural residues to be functional (Fig. S1).

The well-known conservative nature of Type III PKSs was utilized to identify the catalytically important residues in \textit{RaALS} using ClustalW multiple sequence alignment tool (Fig. 3). Importantly, the highly conserved amino acid residues found in nearly all Type III PKSs were also observed to maintain their identity in the primary amino acid sequences of \textit{RaALS}. The multiple sequence alignment revealed that the active-site residues in CHS family viz Cys\textsuperscript{165}, Phe\textsuperscript{216}, His\textsuperscript{304} and Asn\textsuperscript{337} (marked *+) are well conserved in \textit{RaALS}. The characterized enzyme also displayed conserved active-site residues, which include Phe\textsuperscript{139}, Gly\textsuperscript{164}, Gly\textsuperscript{168}, Leu\textsuperscript{215}, Asp\textsuperscript{218}, Gly\textsuperscript{264}, Gly\textsuperscript{336}, Phe\textsuperscript{305}, Gly\textsuperscript{306}, Gly\textsuperscript{307} (marked *+)}
along with ‘GVLFGF’ (displayed by green block) signature motif as found in CHSs. In the malonyl-CoA motif, Glu315, Lys317, Leu318, Leu320, Glu321, Lys323 and Arg330 (displayed by red block) remained conserved in \(RaALS\). Additionally, five out of the seven amino acid residues including Thr133, Met138, Phe216, Phe267 and Pro376 (marked by ‘+’) which are involved in formation of the cyclization pocket (part of the active-site architecture of the enzyme) were found to be conserved in \(RaALS\) with an exception of Ile256 and Gly258 which were non-synonymously changed to Met and Leu, respectively. As discussed in our previous investigation (Pandith et al. 2016), the cyclization pockets’ topology

---

**Fig. 3** Multiple sequence alignment; Sequence alignment of \(RaALS\) with chalcone synthase proteins from *Medicago sativa*, *Gerbera hybrid cultivar*, *Cannabis sativa*, *Fragaria x ananassa* and *Aloe arborescens*. Aligned residues are highlighted from black to grey to white background according to their identity. The catalytic triad residues are marked with ‘*’, ‘$’ sign shows active-site residues. ‘+’ sign displays amino acid cyclization pocket residues whereas ‘S’ signifies the variation observed in \(RaALS\) with respect to other sequences. The red block surrounds the malonyl-CoA-binding motif and green block shows signature motif in CHS.

---
guides polyketide folding and thereby product formation. Apparently, any change in the cyclization pocket may shake the stereochemistry of the cyclization reaction and temper with product selectivity. Though subject to further empirical experimental findings, it might be one of the possible reasons for difference in substrate selectivity and product formation for \textit{RaALS} which derails a bit from well-studied Type III PKSs, CHSs, in particular. Furthermore, three of the inert residues (Thr$^{197}$, Gly$^{256}$, and Ser$^{338}$; numbering in \textit{Medicago sativa} CHS) affecting the shape/volume of the initiation/elongation cavity of active site are known to hold substantial importance in controlling the growing polyketide product (Ferrer et al. 1999; Lim et al. 2016; Songsiriritthigu et al. 2020). These residues were found to be non-synonymously replaced in \textit{RaALS} with Ala, Leu, and Thr, respectively. This event is also anticipated to play a possible role in altered catalytic activity of the isolated enzyme.

Pertinently, there are no crystal structures of aloesone synthase present in Protein Data Bank (PDB). Therefore, homology model of \textit{RaALS} was built using SWISS-MODEL online web service. The crystal structure of CHS with PDB ID 4YJY (Go et al. 2015) was chosen as template structure by BLASTp search against PDB database. Template structure showed 64% sequence identity and 99% query coverage against the query sequence of \textit{RaALS}. The homology model generated by Swiss model is shown in Fig. 4a. Moreover, \textit{RaALS} shares same evolutionary family with CHS and CHS-like Type III PKSs. Indeed, homology modeling of \textit{RaALS} shows five-layer αβαβαβ core structure similar to that of CHSs (Weng and Noel 2012). The RMSD between template structure and model structure is 0.11 Å which supports the high similarity between both structures and suggests a good model quality. Additionally, the model was validated using Ramachandran plot predicted by PROCHECK. The plot showed 94.1% of the residues in the core region, 5.3% in allowed region while 0.3% in disallowed region as shown in Table S3 and Fig. 4b. Residues in the disallowed region were named and colored red on the plot. This revealed that majority of amino acids were in phi–psi distribution of Ramachandran plot; hence, the model is of good quality with overall residue–residue geometry and good stereo-chemical quality. Furthermore, the ProSA analysis of model was computed to understand erroneous energies of each amino acid in the protein (Fig. 4c). As per the plot, all the amino acids show negative energies. The model was found to be ideal with regions mostly falling in correctly determined region by QMEAN Plot (Fig. 4d). The verified 3D score of model was found to be 99.74% (Table S3, Fig. 4e) representing very high agreement of 3D structure to the 1D sequence. Hence, the model proved to be validated in terms of structural geometry and energy profiles suggestive of the perfect starting point for prospective phases of model analyses.

Phylogenetic tree was constructed to understand the evolutionary relationship of \textit{RaALS} (Fig. 5) with orthologous Type III PKSs. Deduced protein sequence from cloned \textit{RaALS} gene was employed to construct the phylogenetic tree using Maximum likelihood Method with MEGA X software (Kumar et al. 2018). CHS from Chinese peony (\textit{Paeonia lactiflora}) was used as an outgroup.
to root the phylogenetic tree. The taxon is a member of the family Paeoniaceae (order, Saxifragales) that is known to share a close evolutionary association with Caryophyllales. The results obtained here indicated that the Type III PKSs from Caryophyllales show an interesting evolutionary history. RaALS from R. australe is seen grouped with PKS enzymes from members of Polygonaceae and Plumbaginaceae both of which belong to the Caryophyllales order. In general, most of the PKSs from genus Rheum form a separate monophyletic group like the ones from Polygonum, Fallopia, Fagopyrum and Dianthus. The specific heptaketide producing PKS characterised in the present study exhibited a close relationship with its orthologous member from the congeneric species R. palmatum (Fig. 5). This little clad advocates their early divergence. In other words, the two members seem to have diverged before the speciation event as shown for R. australe CHS members in our previous investigation (Pandith et al. 2016).

**Molecular docking of RaALS**

Acetyl-CoA is the starter molecule for RaALS and carries out successive six condensation reactions combined with malonyl-CoA to produce aloesone (Fig. 1). To understand binding properties of acetyl-CoA and aloesone with aloesone synthase, docking studies were performed. Certainly, molecular docking is the powerful tool to predict binding modes of ligand with protein structure. Docking studies of acetyl-CoA and aloesone helped to gain insights into interactions at molecular level. Acetyl-CoA being a large bulky molecule takes a large cavity for binding. Further, the docked complexes were visualized by PyMOL to observe interactions of protein with substrate and the product. We found that acetyl-CoA shows major interactions with Leu257 and Thr339 active-site residues, and Ala198 (Fig. 6a). Ala198 is reported to help in extension of the buried pocket, while Leu257 and Thr339 serve as the catalytically important residues (Abe et al. 2006). Interestingly, all the docked complexes of acetyl-CoA and RaALS show high binding energy which might be due to bulky size of molecule or possibly the limitation of docking algorithm to dock the large ligand. On the other hand, product aloesone was observed to bind deeper inside the pocket with binding energy of ~ 6.86 kcal/mol suggesting better binding (Fig. 6b). Surprisingly, aloesone was found to show no interactions with active-site residues which might help in its easy release from the protein. Moreover, aloesone synthase is found to go through conformational changes in binding site to perform polyketide chain elongations (Abe et al. 2006) which might be the possible reason for no interactions with aloesone.

**Recombinant heterologous expression in alternate microbial hosts**

The cds region of RaALS was expressed in E. coli BL21 (DE3) to establish the identity of isolated PKS. The recombinant protein was expressed under the control of Ptac hybrid-promoter using pGEX-4 T-2, an E. coli expression vector induced by IPTG. SDS-PAGE analysis demonstrated that the fusion protein showed optimum expression when induced with 0.8 mM IPTG for 6 h at 30 °C (Fig. 7). The molecular weight of the expressed recombinant enzyme was apparently ~69 kDa and was in agreement with the predicted molecular mass of RaALS (calculated from its deduced amino acid sequence) including that of GST (25.99 kDa). In our previous investigation (Pandith et al. 2016), we demonstrated the functional expression of recombinant Type III PKSs in soluble fraction with a homogeneous molecular weight of approximately 42 kDa, excluding that of the associated GST tag. But, the target gene in present study could not be expressed in soluble fraction in sufficient amount as most part of it remained localized to the inclusion bodies hampering its purification. Nonetheless, RaALS was expressed and intended to be purified for in vitro functional validation. Certainly, in vitro protein characterization by enzymatic, spectrophotometric or calorimetric methods would have further substantiated its function by validating the substrate specificity, pH, multi-functionality and highlighted the enzyme kinetic parameters including maximum velocity and rate of reaction. However, lack of purified protein impeded protein characterization. Furthermore, switching of the host system from prokaryotic to eukaryotic (S. pombe) organism to purify the target Type III PKS using pDS472a expression vector also failed to yield purified protein due to abstruse reasons.

**Determination of gene copy number—RaALS exists as single copy in R. australe**

The event of gene duplication has occurred several times over the course of evolution, and it is estimated that approximately 65% of plant genes are duplicated which in turn sway the overall architecture and function of associated genomes (Panchy et al. 2016). Though with a largely conserved gene structure, the CHS and CHS-like (includes RaALS as well) gene family also appears to have one to many members exhibiting novel or promiscuous functions. For instance, in bread wheat (Triticum aestivum L.), 10 copies of CHS gene have been identified (Glagoleva et al. 2019) and 14 members in the paleopolyploid soybean (Glycine max [L] Merr) (Vadivel et al. 2018) highlighting their importance in the synthesis of secondary chemical constituents viz. (iso)flavonoids and their derived phytoalexins. Similarly, a total of 7 CHS gene family members have been found each in the eggplant.
(Solanum melongena L.) (Wu et al. 2020) and in Petunia hybrida (Koes et al. 1987) while 6 in turnip (Brassica rapa) that encode light-responsive redundant genes exhibiting varied expression in different tissues for metabolite (flavonoid) biosynthesis (Zhou et al. 2013). Now, to estimate the number of RaALS gene copies in R. australe and to validate this Type III plant PKS, genomic southern blot analysis was performed under high stringency conditions as discussed in some of our previous studies (Rana et al. 2013; Dhar et al. 2014; Pandith et al. 2016). Three restriction enzymes used viz BamHI and EcoRI—without any restriction site in the cds region of the gene, and NotI—with a single cutting site, generated the restriction pattern which was consistent with a single-copy gene (Fig. 8). A single band was scored in BamHI and EcoRI digested DNA samples and two bands were observed with NotI digestion which indicated that PCR synthesized DIG-labeled probe hybridized to the specific gene of interest. The results obtained in southern hybridization thus suggest that R. australe genome contains a single copy of RaALS gene. In fact, the gene family members vary in number from one to the other species owing to the need of attributes they are ascribed to. Interestingly, including two earlier characterized members (Pandith et al. 2016), so far, we have identified three members of the Type III PKSs from R. australe which exhibit varied/promiscuous functions in the path of secondary metabolism vis-à-vis dynamic eco-physiological attributes besides environmental agitations of varied nature. Though subject to further empirical experimental investigations, the Type III PKS gene family does not seem to have evolved much (to generate its various homologous members) in this important medicinal herb over the course of evolution.

RaALS is highly expressed in leaves of R. australe

The Type III plant PKSs, in general, exhibit differential expression pattern in different tissues and in response to the environmental perturbations. For instance, in a recent study on three different species of chili pepper (Capsicum annuum, C. chinense, and C. baccatum), 13–14 identified PKSs were shown to display varied expression consistent with the maturation of the fruit vis-à-vis metabolite (flavonoid) accumulation trends (Kan et al. 2020). With an endeavor to characterize the North American grape cultivars for the production of anthocyanin antioxidants, Davis et al. (2012) observed variations in the expression of CHS enzymes in the skin and flesh of the berries which were supposed to be due to the ontogenetic stages and environmental perturbations. In our previous study (Pandith et al. 2016) on R. australe CHSs—the simplest representatives of Type III PKSs, a comparative analyses of flavonoids and anthraquinones revealed significantly higher content of rutin (the predominant flavonoid) and chrysophanol (the predominant anthraquinone) in leaves as compared to that of stem

---

**Fig. 5** Evolutionary relatedness of orthologous Type III plant PKSs: the evolutionary history was inferred using the Maximum Likelihood method and JTT matrix-based model. The tree with the highest log likelihood (−7803.79) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The numbers on the nodes indicate the bootstrap values after 500 replicates. The bar indicates an evolutionary distance of 0.2%. This analysis involved 54 amino acid sequences. Evolutionary analyses were conducted in MEGA X.

**Fig. 6** Molecular Docking of RaALS; (a) Starter molecule Acetyl-CoA. (b) product aloesone. Protein is rendered in cartoon coloured green while ligand is shown in licorice in magenta colour. H-bond interactions are shown with dotted black line.
and the root tissues. The same samples were analysed using relative quantitative real-time PCR to examine the spatial regulation of \( \text{RaALS} \) gene in different tissues (leaf, stem and root) of \( \text{R. australe} \) (Fig. 9a). Though expressed in all tissues, the transcripts of \( \text{RaALS} \) gene exhibited a distinct expression pattern with leaves \((1.01 \pm 0.17)\) showing maximum transcript abundance, followed by stem \((0.85 \pm 0.63)\) and the root \((0.19 \pm 0.01)\). Here, \( \text{RaALS} \) expression corroborates well with the specific metabolite accumulation we had observed in leaves as compared to the other plant tissues (Pandith et al. 2016); and supplemented by another study on \( \text{Fagopyrum esculentum} \), a Polygonaceae member well studied for rutin biosynthesis (Li et al. 2010). Similarly, and like in some grass species with high contents of UV-B protecting

`RaALS exhibits differential expression pattern along the altitudinal gradient`

Altitude, one of the significant natural experiments to test the ecological and evolutionary responses of living world to the non-living geo-physical influences, is known to have a prominent effect on accumulation of the secondary chemical constituents in plants. Indeed, among other ecological factors like temperature, light exposure, etc., altitude is known to design the content and composition of plants vis-à-vis secondary chemical constituents (Medda et al. 2021). Moreover, a recent investigation projected the role of altitude as an operational factor on the synthesis and accumulation of phenolics in strawberry fruits in relation to their antioxidant activity (Guerrero-Chavez et al. 2015). Similarly, and like in some grass species with high contents of UV-B protecting

Fig. 8 Southern blot analysis of \( \text{RaALS} \): The genomic DNA (>20 μg) isolated from \( \text{Rheum australe} \) was digested with BamHI, EcoRI (both non-cutters; do not cut anywhere in the reading frame of the \( \text{RaALS} \) gene) and NotI (single-cutter; has a single nicking site in the \( \text{RaALS} \) ORF) restriction enzymes. The restriction-digested samples were separated on 0.8% agarose gel (loaded in alternate wells of the gel to avoid ambiguity), blotted onto a nylon membrane and hybridized with DIG-labeled ORFs of \( \text{RaALS} \) as probe. The BamHI and EcoRI digested samples yielded a single band (as they did not cut the ORF anywhere), whereas the single-cutter NotI gave two bands (\( \text{RaALS} \) cut at single position [at position 659, 5’ extension] by the NotI enzyme to give two fragments both of which were identified and thereby hybridized by the gene-specific probe)
flavonoids (luteolin and orientin), high-altitude maize plants accumulate certain flavones in leaves and silks to prevent themselves from deleterious UV-B effects (Falcone Ferreyra et al. 2012). Pertinently, in our previous investigations, we found a general increasing trend of secondary metabolite constituents with increasing altitude (Pandith et al. 2014, 2016), and the ploidy status as well (Farooq et al. 2013). In particular, major flavonoid constituents viz. naringenin and rutin were seen to accumulate at greater concentrations at higher altitudes. This observation prompted us to ascertain the levels of \textit{RaALS} transcripts in same tissues (as used in previous investigations) collected from four different geographic regions, and to correlate the findings with the targeted metabolite accumulation. Fortunately, we noticed consonance in \textit{RaALS} gene promoter. The putative core promoter consensus sequences and the motifs with significant similarity to the previously identified \textit{cis}-acting elements are highlighted and the names are given. TIS represents the transcription initiation site and 5' untranslated region (UTR) is depicted with an arrow. \textbf{(d-f)} Time course expression characteristics of \textit{RaALS} gene in micro-propagated \textit{R. australe} in response to the elicitations by 0.1 mM MeJ (d), 0.1 mM SA (e), and 1500 µJ/m² UV-B radiation (f). The in vitro raised cultures were precultured in MS liquid medium for about 2 w, elicited and further harvested after different time intervals and the concentration of MeJ and SA for plant treatments was kept as 0.1 mM.

Identification of putative \textit{cis}-regulatory elements from isolated promoter sequence

In eukaryotes, the process of transcription typically follows an event of specific recognition between a transcription
factor (TF) and its binding site (cis-element) which is usually located at the upstream region of a gene. Knowing the particular sequence of this TF-binding site holds critical importance in gene regulation (Walhout 2006). Taking this in mind, and to elucidate the transcriptional regulation of RaALS, genome walking approach was used to isolate the 5 upstream region of this gene. This strategy led to the isolation of 416 bp promoter sequence which was scanned in silico with PlantCare (Lescot et al. 2002) and PLACE (Higo et al. 1999) online tools for the identification of various putative cis-regulatory elements. Pertinently, the RaALS promoter sequence was found to contain various important cis-regulatory elements (Fig. 9c, Table 1), besides having a typically high A+T content (56.25%) usually seen in other plant promoters.

The predicted transcription initiation site (TIS, +1) was found to be located at 116 bp upstream of the RaALS start codon. SOGO (New PLACE) analyses of the isolated promoter sequence showed the location of putative TATA box (essential for initiating transcription process) at position 98 (+) upstream with respect to TIS. The eukaryotic promoter regulatory consensus cis-acting element CAAT box, commonly found in promoter and enhancer regions, was observed at positions 153 (+), 220 (+), 261 (−), 262 (−), 380 (−), 381 (−). The light-responsive elements viz. Box I, GATA box and GT-1 motif were also found in the RaALS promoter region. Further, several other functional elements, such as MYB core, CGTCA motif, TC-rich repeats, TGACG motif/TCA element and Skn-1 motif, were also found at various positions in the promoter sequence. These elements are reportedly known for various activities including defense and stress responsiveness and tissue-specific expression, etc. MYB elements, found in numerous plant promoters, are known to serve as binding sites for R2R3 MYB TFs (Duraisamy et al. 2018). Additionally, the light-responsive element, GT-1 motif, is also known to be involved in SA elicited gene expression (Bhat et al. 2014b). Specific regulatory elements (Box I, CGTCA motif, TGACG motif/TCA element, TC-rich repeats and GT-1 motif) from RaALS promoter region were selected to determine their role in stress responsiveness under in vitro conditions to examine their inducible/repressible nature.

Expression of RaALS is inducible by MeJA, SA and UV-B elicitions

Elicitors, in general, are the chemicals or bio-factors from varied sources which can induce the physiological and morphological changes within a living host. In plants, the elicitors, biotic or abiotic, generate a range of defense reactions which mostly include the accumulation of an array of defensive secondary chemical constituents. Indeed, contemporary developments in comprehension of plant signaling pathways have flagged an approach of utilizing elicitor-induced secondary metabolite variation in plants as a projecting device for potential pathway elucidation, regulation and prospective engineering studies. Therefore, to examine and evaluate the constitutive or inducible nature of RaALS gene, and to validate the results generated with the cis-regulatory elements identified from the promoter region of the target gene, RaALS transcripts were assayed in response to the exogenous induction of MeJ (0.1 mM), SA (0.1 mM), and UV-B light (1500 µJ m⁻²) exposure.

Jasmonic acid, its related (MeJ) and even conjugated compounds, is reportedly well-known transducers and mediators leading to the production and accumulation of various classes of plant secondary metabolites (Tamogami et al. 1997; Farmer et al. 2003; Zhao et al. 2005). In present study, MeJ treatment significantly instigated the expression of RaALS gene with a steep increase. The mRNA levels peaked at 24 h with nearly 18-fold increase as compared to that of control. However, the transcript level slightly declined afterwards as shown in the Fig. 9d. These observations are in conformity with some of the earlier investigations on Type III plant PKSs (Richard et al. 2000; Yu et al. 2015; Pandith et al. 2016).

Though not a universal agent to induce production of the defensive compounds in plants, SA is known to tempt expression of genes related to the generation of some classes of plant secondary metabolites (Taguchi et al. 2001). In R. australe, SA elicited a gradual, but nearly significant, increase in the transcript abundance of RaALS gene. About two-fold rise was registered in the expression profile at 24 h post induction after which the expression declined to less than the control value (Fig. 9e).

The varied response of plants against irradiation acts by triggering protection/repair mechanisms which mostly involve biosynthesis of the UV-absorbing secondary metabolite constituents; in particular, phenolic compounds (Hahlbrock and Scheel 1989; Dao et al. 2011). Earlier studies have revealed that these UV-absorbing chemical compounds accumulate distinctly upon UV-B (280–320 nm) exposure which has been further shown to be the consequence of upsurge in transcript levels of various enzymes of phenylpropanoid biosynthetic pathway including Type III PKSs (Davies and Schwinn 2003; Jaakola et al. 2004). To analyze the effects of UV-B irradiation on RaALS transcript abundance, the in vitro raised plantlets were subjected to UV-B light treatment for 9 h under controlled conditions. As shown in the Fig. 9f, the transcripts of RaALS gene displayed an irregular, but increasing, trend in their accumulation. Nearly a 6- and fourfold increase was observed at 9 and 3 h post UV-B induction, respectively. In general, the results obtained were in agreement with some of the earlier investigations.
of Type III PKS gene family members on Soybean (Tuteja et al. 2004) and Mulberry (Li et al. 2016).

Conclusion and future perspectives

Owing to the treasury of polyketide-derived bio-active secondary chemical constituents, PKSs have cultivated a growing and vital interest among scientists toward the pol- yketide biosynthetic machinery. Indeed, the simplicity of Type III PKSs has made them soft and achievable targets for their understanding and rational protein engineering. In this context, the present work is an endeavor to understand the molecular basis of biosynthesis of species-specific bio-active phyto-constituents. Pertinently, we have successfully isolated and characterized the full-length RaALS gene in the current study. Bioinformatics-based structural elucidation and binding analysis enlighten its molecular mechanisms of catalysis. The copy number analysis using southern blotting proved RaALS to be single-copy gene. Additionally, promoter isolation and analysis provided insights into expression regulation of this CHS-like family member, RaALS. Certainly, the transcript levels of RaALS differed spatially and with altitude, and were also found to be modulated by various cues (exogenous elicitations) based on the identified cis-regulatory elements vis-à-vis accumulation of major bio-active chemical constituents. Further, the promiscuous behaviour of RaALS in generating a heptaketide intermediate was also highlighted in relation to the amino acid changes in some of the vital (and conserved) positions of this peculiar enzyme. Present work may pave the way toward understanding the role of CHS-like (and other) Type III PKSs in the biosynthesis of pharmaco logically important phytoconstituents; anthraquinones, in particular. Moreover, the recent advancements in synthetic and systems biology have opened up several new ways toward understanding and realizing the potential of PKSs and further utilizing them, through rational enginee ring of their biosynthetic machinery, for the generation of more diverse and novel lead compounds with desirable pharmacological efficacy.

Supplementary Information  The online version contains supplementary material available at https://doi.org/10.1007/s11816-022-00741-5.

Author contributions  Conceived and designed the experiments: SAP, SL; performed the experiments: SAP, SR, SB, ND; analyzed the data: SAP, VB, ND, MAS, SL; contributed reagents/materials/analysis tools: SAP, SL; wrote the paper: SAP, SR, VB, ND.

Funding  Work in the SAP laboratory is supported by Department of Science and Technology (DST), Govt. of India, under the INSPIRE Faculty Scheme [DST/INSPIRE/04/2016/0001059]. Earlier, the work was partly supported by financial grant from Council of Scientific and Industrial Research (CSIR)-Indian Institute of Integrative Medicine under Major Lab Project MLP-3012.

Declarations

Conflict of interest  The authors declare that there is no conflict of interest.

References

Abe I, Watanabe T, Lou W, Noguchi H (2006) Active site residues governing substrate selectivity and polyketide chain length in aloesone synthase. FEBS J 273(1):208–218
Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38(suppl 2):W529–W533
Awashti P, Mahajan V, Jamwal VL, Kapoor N, Rasool S, Bedi YS, Gandhi SG (2016) Cloning and expression analysis of chalcone synthase gene from Coleus forskohlii J. Genetics 95(3):647–657
Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27(3):343–350
Bhat WW, Dhar N, Razdan S, Rana S, Mehra R, Nargotra A, Dhar RS, Ashraf N, Vishwakarma R, Lattoo SK (2013) Molecular characterization of UGT94F2 and UGT86C4, two glycosyltransf erases from Picrorhiza kurrooa: comparative structural insight and evaluation of substrate recognition. PLoS ONE 8(9):e73804
Bhat WW, Rana S, Dhar N, Razdan S, Pandith SA, Vishwakarma R, Lattoo SK (2014a) An inducible NADPH–cytochrome P450 reductase from Picrorhiza kurrooa—an imperative redox partner of cytochrome P450 enzymes. Funct Integr Genomics 14(2):381–399
Bhat WW, Razdan S, Rana S, Dhar N, Wani TA, Qazi P, Vishwakarma R, Lattoo SK (2014b) A phenylalanine ammonia-lyase ortholog (PkPAL1) from Picrorhiza kurrooa Royle ex. Benth: molecular cloning, promoter analysis and response to biotic and abiotic elicitors. Gene 547(2):245–256
(2009) CTAB DNA extraction buffer. Cold Spring Harb Protoc 2009(3):pdb.rec11718. https://doi.org/10.1101/pdb.rec11718
Dao T, Linthorst H, Verpoorte R (2011) Chalcone synthase and its functions in plant resistance. Phytochem Rev 10(3):397
Davies KM, Schwinn KE (2003) Transcriptional regulation of secondary metabolism. Functional Plant Biol 30(9):913–925
Davis G, Ananga A, Krastanova S, Sutton S, Ochieng JW, Leong S, Davies KM, Schwinn KE (2003) Transcriptional regulation of secondary metabolism. Functional Plant Biol 30(9):913–925
Davies G, Ananga A, Krastanova S, Sutton S, Ochieng JW, Leong S, Tsoleva V (2012) Elevated gene expression in chalcone synthase enzyme suggests an increased production of flavonoids in skin and synchronized red cell cultures of North American native grape berries. DNA Cell Biol 31(6):939–945
Deng X, Bashandy H, Ainasoja M, Kontturi J, Pietiäinen M, Laitinen RS, Ashkenazy H, Erez E, Martz E, Pupko T, Lattoo SK (2014) Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of Gerbera hybrida. New Phytol 201(4):1469–1483
Dhar N, Rana S, Razdan S, Bhat WW, Hussain A, Dhar RS, Vaishnavi S, Hamid A, Vishwakarma R, Lattoo SK (2014) Cloning and functional characterization of three branch point oxidosqualene cyclases from Withania somnifera (L.) dunal. J Biol Chem 289(24):17249–17267
Duraisamy G, Mishra A, Kocábek T, Matoušek J (2018) Activation of polyketide synthase gene promoter in Cannabis sativa by...
heterologous transcription factors derived from *Hamulus lupulus*. Biol Plant 62(2):250–260

Falcone Ferreyra ML, Rius S, Catas P (2012) Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front Plant Sci 3(222):1–15

Farmer EE, Almáras E, Krishnamurthy V (2003) Jasmonates and related oxylipins in plant responses to pathogenesis and herbivory. Curr Opin Plant Biol 6(4):372–378

Farooq U, Pandith SA, Saggoo MIS, Lattoo SK (2013) Altitudinal variability in anthraquinone constituents from novel cytotypes of *Rumex nepalensis* Spreng—a high value medicinal herb of North Western Himalayas. Ind Crop Prod 50:112–117

Ferrer J-L, Jez JM, Bowman ME, Dixon RA, Noel JP (1999) Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat Struct Biol 6(8):775–784

Flores-Sanchez IJ, Verpoorte R (2009) Plant polyketide synthases: a fascinating group of enzymes. Plant Physiol Biochem 47(3):167–174

Funabashi M, Funa N, Horinouchi S (2008) Phenolic lipids synthesized by type III polyketide synthase confer penicillin resistance on *Streptomyces griseus*. J Biol Chem 283(20):13983–13991

Glagoleva AY, Ivanisenko NV, Khlestkina EK (2019) Organization and evolution of the chalcone synthase gene family in bread wheat and related species. BMC Genet 20(1):25–33

Go MK, Wongstachtsorn J, Cheung VWN, Chow JY, Robinson RC, Yew WS (2015) Synthetic polyketide enzymology: platform for biosynthesis of antimicrobial polyketides. ACS Catal 5(7):4033–4042

Guerrero-Chavez G, Scampicchio M, Andreotti C (2015) Influence of the site altitude on strawberry phenolic composition and quality. Sci Hortic 192:1–18

Guo X-D, Ma Y-J, Parry J, Gao J-M, Yu L-L, Wang M (2011) Phenolics content and antioxidant activity of tartary buckwheat from different locations. Molecules 16(12):9850–9867

Hahlbrock K, Scheel D (1989) Physiology and molecular biology of flavonoid synthesis. Annu Rev Plant Biol 40(1):347–369

Hertweck C (2009) The biosynthetic logic of polyketide diversity. Angew Chem Int Ed 48(26):4688–4716

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

Jaakola L, Määttä-Riihinen K, Kärenlampi S, Hohtola A (2004) Activation of flavonoid biosynthesis by solar radiation in bilberry (*Vaccinium myrtillus* L.). Planta 218(5):721–728

Jeelani SM, Bhat WW, Dhar N, Razdan S, Wani TA, Shah MA, Bedi Y (2014) Evaluation of anthraquinones from Himalayan rhubarb (*Rheum emodi* Wall. ex Meissn.) as antiproliferative agents. S Afr J Bot 95:1–8

Koes RE, Spitaler R, Schlorhaufer PD, Ellmerer EP, Merfort I, Bortenschlager S, Stuppern H, Zidorn C (2006) Altitudinal variation of second-generation polyphenols in *Garcinia mangostana* L. pericarps reveals the basis for substrate specificity and catalysis. Acta Crystallogr Sect F 176(12):597–603

Li X, Park NI, Xu H, Woo S-H, Park CH, Park SU (2010) Differential expression of flavonoid biosynthesis genes and accumulation of phenolic compounds in common buckwheat (*Fagopyrum esculentum*). J Agr Food Chem 58(23):12176–12181

Li H, Liang J, Chen H, Ding G, Ma B, He N (2016) Evolutionary and functional analysis of mulberry type III polyketide synthases. BMC Genom 17(1):540

Lim YP, Go MK, Yew WS (2016) Exploiting the biosynthetic potential of type III polyketide synthases. Molecules 21(6):806

Medda S, Sanchez-Ballesta MT, Romero I, Dessena L, Mulas M (2021) Expression of structural flavonoid biosynthesis genes in dark-Blue and White Myrtle Berries (*Myrtus communis* L.). Plants 10(2):316

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30(16):2785–2791

Nair DR, Anand S, Verma P, Mohanty D, Gokhale RS (2012) Genetic, biosynthetic and functional versatility of polyketide synthases. Curr Sci 277–287

Panchy N, Lehti-Shiu M, Shiu SH (2016) Evolution of gene duplication in plants. Plant Physiol 171(4):2294–2316

Pandith SA, Hussain A, Bhat WW, Dhar N, Qazi AK, Rana S, Razdan S, Wani TA, Shah MA, Bedi Y (2014) Evaluation of anthraquinones from Himalayan rhubarb (*Rheum emodi* Wall. ex Meissn.) as antiproliferative agents. S Afr J Bot 95:1–8

Pandith SA, Dhar N, Rana S, Bhat WW, Kushwaha M, Gupta AP, Shah MA, Vishwakarma L, Lattoo SK (2016) Functional promiscuity of two divergent paralogs of type III plant polyketide synthases. Plant Physiol 171(4):2599–2619

Pandith SA, Dar RA, Lattoo SK, Shah MA, Reshi ZA (2018) *Rheum australe*, an endangered high-value medicinal herb of North Western Himalayas: a review of its botany, ethnomedical uses, phytochemistry and pharmacology. Phytochem Rev 17(3):573–609

Pandith SA, Ramazan S, Khan MI, Reshi ZA, Shah MA (2020) Chalcone synthases (CHSs): the symbolic type III polyketide synthases. Planta 251(1):15

Petteresen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25(13):1605–1612

Rana S, Lattoo SK, Dhar N, Razdan S, Bhat WW, Dhar RS, Vishwakarma R (2013) NADPH-cytochrome P450 reductase: molecular cloning and functional characterization of two paralogs from *Withania somnifera* (L.) Dunal. PLoS ONE 8(2):e57068

Richard S, Lapointe G, Rutledge RG, Séguin A (2000) Induction of chalcone synthase expression in white spruce by wounding and jasmonate. Plant Cell Physiol 41(8):982–987

Sayers EW, Beck J, Bolten EE, Bourieux D, Brister JR, Canese K, Chen H, Church GM, Clements J, Davis NW, Eppig JT, Lewis S, Maglott D,Mesirov JP, Mi H, Mungall C, Nickerson S, O’Donoghue S, O’Meara E, Riesch R, Rojas R, Schuler GL, St内外. Sayers EW, Beck J, Bolton EE, Courchesne D, Funk K, Kim S, Klimek W, Marchler-Bauer A (2021) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 49(D1):D10 (Database issue)

Sievers F, Higgins DG (2021) The clustal omega multiple alignment package. In: Kato I (ed) Multiple sequence alignment. Humana, New York, pp 3–16

Songsiriritthigul C, Nualkaew N, Ketudat-Cairns J, Chen CJ (2020) The crystal structure of benzophenone synthase from *Garcinia mangostana* L. pericarps reveals the basis for substrate specificity and catalysis. Acta Crystallogr Sect F 176(12):597–603

Spitaler R, Schlorhauder PD, Ellmerer EP, Merfort I, Bortenschlager S, Stuppern H, Zidorn C (2006) Altitudinal variation of secondary metabolite profiles in flowering heads of *Arricua montana* cv. ARBO Phytochem 67(4):409–417
Taguchi G, Yazawa T, Hayashida N, Okazaki M (2001) Molecular cloning and heterologous expression of novel glucosyltransferases from tobacco cultured cells that have broad substrate specificity and are induced by salicylic acid and auxin. Eur J Biochem 268(14):4086–4094

Tamogami S, Rakwal R, Kodama O (1997) Phytoalexin production elicited by exogenously applied jasmonic acid in rice leaves (Oryza sativa L.) is under the control of cytokinins and ascorbic acid. FEBS Lett 412(1):61–64

Tuteja JH, Clough SJ, Chan W-C, Vodkin LO (2004) Tissue-specific gene silencing mediated by a naturally occurring chalcone synthase gene cluster in Glycine max. Plant Cell 16(4):819–835

Vadivel AK, Krysiak K, Tian G, Dhaubhadel S (2018) Genome-wide identification and localization of chalcone synthase family in soybean (Glycine max [L.] Merr). BMC Plant Biol 18(1):1–3

Walhout AJ (2006) Unraveling transcription regulatory networks by protein–DNA and protein–protein interaction mapping. Genome Res 16(12):1445–1454

Wani TA, Pandith SA, Gupta AP, Chandra S, Sharma N, Lattoo SK (2017) Molecular and functional characterization of two isoforms of chalcone synthase and their expression analysis in relation to flavonoid constituents in Grewia asiatica L. PLoS ONE 12(6):e0179155

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienry R, Heer FT, de Beer TAP, Rempfer C, Bordoli L (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46(W1):W296–W303

Weng J-K, Noel JP (2012) Structure–function analyses of plant type III polyketide synthases. Method Enzymol 515:317–335

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407–W410 (suppl 2)

Wu X, Zhang S, Liu X, Shang J, Zhang A, Zhu Z, Zha D (2020) Chalcone synthase (CHS) family members analysis from eggplant (Solanum melongena L.) in the flavonoid biosynthetic pathway and expression patterns in response to heat stress. PLoS ONE 15(4):e0226537

Yu H-N, Wang L, Sun B, Gao S, Cheng A-X, Lou H-X (2015) Functional characterization of a chalcone synthase from the liverwort Plagiochasma appendiculatum. Plant Cell Rep 34(2):233–245

Zeng J, Decker R, Zhan J (2012) Biochemical characterization of a type III polyketide biosynthetic gene cluster from Streptomyces toxytricini. Appl Biochem Biotechnol 166(4):1020–1033

Zhou B, Wang Y, Zhan Y, Li Y, Kawabata S (2013) Chalcone synthase family genes have redundant roles in anthocyanin biosynthesis and in response to blue/UV-A light in turnip (Brassica rapa; Brassicaceae). Am J Bot 100(12):2458–2467

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.