Functional Promiscuity of Two Divergent Paralogs of Type III Plant Polyketide Synthases[1]

Shahzad A. Pandith, Niha Dhar[2], Satiander Rana[3], Wajid Waheed Bhat[4], Manoj Kushwaha, Ajai P. Gupta, Manzoor A. Shah, Ram Vishwakarma, and Surrinder K. Lattoo*

Plant Biotechnology Division (S.A.P., N.D., S.R., W.W.B., S.K.L.), Quality Control and Quality Assurance Division (M.K., A.P.G.), and Medicinal Chemistry Division (R.V.), CSIR-Indian Institute of Integrative Medicine, Jammu Tawi 180001, India; and Department of Botany, University of Kashmir, Srinagar 190006, Jammu and Kashmir, India (M.A.S.)

ORCID IDs: 0000-0001-9197-8378 (S.R.); 0000-0002-8902-3713 (S.K.L.).

Plants effectively defend themselves against biotic and abiotic stresses by synthesizing diverse secondary metabolites, including health-protective flavonoids. These display incredible chemical diversity and ubiquitous occurrence and confer impeccable biological and agricultural applications. Chalcone synthase (CHS), a type III plant polyketide synthase, is critical for flavonoid biosynthesis. It catalyzes acyl-coenzyme A thioesters to synthesize naringenin chalcone through a polyketidic intermediate. The functional divergence among the evolutionarily generated members of a gene family is pivotal in driving the chemical diversity. Against this backdrop, this study was aimed to functionally characterize members of the CHS gene family from *Rheum emodi*, an endangered and endemic high-altitude medicinal herb of northwestern Himalayas. Two full-length cDNAs (1,179 bp each), *ReCHS1* and *ReCHS2*, encoding unique paralogs were isolated and characterized. Heterologous expression and purification in *Escherichia coli*, bottom-up proteomic characterization, high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry analysis, and enzyme kinetic studies using five different substrates confirmed their catalytic potential. Phylogenetic analysis revealed the existence of higher synonymous mutations in the intronless divergents of ReCHS. ReCHS2 displayed significant enzymatic efficiency (V_{max}/K_m) with different substrates. There were significant spatial and altitudinal variations in messenger RNA transcript levels of ReCHSs correlating positively with metabolite accumulation. Furthermore, the elicitations in the form of methyl jasmonate, salicylic acid, ultraviolet B light, and wounding, chosen on the basis of identified cis-regulatory promoter elements, presented considerable differences in the transcript profiles of ReCHSs. Taken together, our results demonstrate differential propensities of CHS paralogs in terms of the accumulation of flavonoids and their relative substrate selectivities.

Plants, as ground-anchored sessile creatures, invest significant amounts of energy in the production of secondary metabolites to combat environmental pressures. These metabolites often are produced through complex and highly regulated biosynthetic pathways under the influence of different enzymatic machineries. One of the important classes of secondary metabolites is phenylpropanoids. These represent a significant proportion of secondary metabolites, encompassing nearly 20% of total carbon in the terrestrial biosphere. Flavonoids exhibit remarkable chemical diversity and ubiquitous occurrence and play an important role in many aspects of plant development, like flower coloration, photoprotection, pollen development, cell wall growth, and response to stress conditions like UV light protection, herbivory, wounding, interaction with soil microbes, and defense against pathogens (Yu and Jez, 2008; Pandey et al., 2015). Apart from performing numerous imperative roles in plants, flavonoids also have been reported as potent phytoceuticals with putative health benefits (Jiang et al., 2015). Flavonols are generally the most copious of all the flavonoids, and

[1] This work was supported by the CSIR-IIIM (Major Lab Project grant no. MLP–3012 and Network Project grant no. BSC–0106).

[2] Present address: Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore, Singapore 117604.

[3] Present address: Genetics, Development and Cell Biology, National Science Foundation Engineering Research Center for Biorenewable Chemicals, Biorenewables Research Laboratory, Iowa State University, Ames, IA 50011.

[4] Present address: Biotransformation, Scion Research New Zealand-Crown Research Institute, 49 Sala Street, Rotorua 3046, New Zealand.

* Address correspondence to sklattoo@iiim.ac.in.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Surrinder K. Lattoo (sklattoo@iiim.ac.in).

S.K.L. and R.V. conceived and designed the experiments; S.A.P. performed the bulk of the experiments as part of his PhD program; experiments and bioinformatics analyses were supported by N.D., S.R., and W.W.B.; support for proteomic and metabolic analyses was provided by M.K. and A.P.G.; data were analyzed by S.A.P., M.S., and S.K.L.; R.V. and S.K.L. contributed reagents/materials/analyses tools; S.A.P. and S.K.L. wrote the article.

www.plantphysiol.org/cgi/doi/10.1104/pp.16.00003

they usually accumulate as glycosides of quercetin, kaempferol, or myricetin in plant vacuoles (Stafford, 1990).

Type III polyketide synthases (PKSs) produce a wide array of natural products, including flavonoids. Chalcone synthase (CHS; EC 2.3.1.74), a well-known and the simplest representative of type III plant PKSs, catalyzes the first committed step in the biosynthesis of flavonoids. It carries out iterative condensation of three acetate units from malonyl-CoA to a favorite and main physiological starter molecule, 4-coumaroyl-CoA, leading to the synthesis of an aromatic tetraketide, naringenin chalcone (Fig. 1). In plants, chalcone is converted to its isoform (2*S*)-5,7,4′-trihydroxyflavanone (naringenin) by an enzyme, chalcone isomerase. The mechanistic dimension of iterative condensation has been facilitated greatly by the elucidation of the x-ray crystal structure of alfalfa (*Medicago sativa*) CHS2 (Ferrer et al., 1999). The crystal structure revealed that CHS is homodimeric (monomer unit, 42–45 kD) and that each monomer has the active site buried deep inside with catalytic triad residues (Cys-164, His-303, and Asn-336) sitting at its top. Topologically, the catalytic triad along with Phe-216 intersect with three interconnected cavities, a CoA-binding tunnel, a coumaroyl-binding pocket, and a cyclization pocket, to form the active site architecture of CHS. The active site consists of a bilobed initiation/elongation/cyclization cavity, with one lobe thought to choose the aromatic moiety of *p*-coumaroyl-CoA and the other housing the growing polyketide chain (Abe and Morita, 2010).

Rheum emodi (syn. *Rheum australe*, Polygonaceae) is a rich repository of pharmaceutically important secondary metabolite constituents like flavonoids, anthraquinones, and stilbenoids. Most of these compounds are produced by various condensation and cyclization events of polyketidic intermediates in the single active site of CHS. The interest in secondary metabolites and the pathways associated with their biosynthesis has seen a great surge with the advent of synthetic biology approaches to harness the production of high-value bioactive molecules in heterologous hosts. Recently, substantial progress has been made using combinatorial biosynthetic approaches to produce high-value bioactive compounds like artemisinic acid (Walter, 2014), opiates (Ehrenberg, 2015), and aglyconic etoposide (Lau and Sattely, 2015) in homologous and/or heterologous hosts. Owing to the wide scope of flavonoids in both plant physiology and human nutrition, a great deal of research has been directed toward their biosynthetic pathway in a range of plant species. It has extended primarily from the model plants to various economically important species, including apple (*Malus domestica*), grape (*Vitis vinifera*), potato (*Solanum tuberosum*), strawberry (*Fragaria* spp.), and tomato (*Solanum lycopersicum*; Deng et al., 2014). However, the limited availability of these beneficial compounds from their natural sources makes flavonoids prime candidates to be synthesized in engineered microbes like *Escherichia coli* and *Saccharomyces cerevisiae* (Luo et al., 2015).

Plant genomes have emerged to direct the process of favoring evolutionary events as a primary means of acquiring biochemical and developmental flexibility. To persist over long periods of evolutionary time, nearly all plant genes are represented within most plant genomes as small to large gene families (Han et al., 2006). For instance, in many angiosperms, CHS is encoded by a multigene family with members ranging from two (*Zea mays*), three (*Gerbera hybrida*), six (*Ipomoea purpurea*), seven (*Pisum sativum*), eight (*Physcomitrella patens*), nine (*Glycine max*), and even 10 in some species like *Petunia hybrida* (Jiang et al., 2006; Deng et al., 2014). These gene families emerge as a result of selective gene duplication and subsequent nucleotide substitution events. The existence of a multigene family may either increase the amount of enzyme/protein product synthesized or help in the evolution of new promiscuous functions. These incidences selectively help the plant to adapt to changing environmental conditions. The dynamics of gene family evolution is well illustrated by CHSs. Although the CHS genes have been found to be structurally conserved, their differential evolutionary rates often are correlated to their promiscuous functions (Pang et al., 2005).

In this study, the approaches of molecular and functional characterization were used to understand the evolutionary implications vis-à-vis the functional versatility of two distantly related members of the CHS gene family isolated from *R. emodi*. The purified and characterized proteins were identified by the bottom-up approach of proteomic analysis and further validated by enzyme kinetics using five different substrates. Site- and tissue-specific expression profiling also was performed and was correlated with the accumulation pattern of major flavonoid and anthraquinone constituents. Moreover, exogenous and endogenous elicitors selected on the basis of isolated cis-acting promoter elements were shown to alter the expression pattern of ReCHSs, which corroborated well the flux in the accumulation of secondary products. Additionally, the flavonoid accumulation in different tissues was further analyzed by staining with a fluorescent dye, diphenylboric acid 2-aminoethyl ester (DPBA). Our results suggest that the ReCHS1 paralog displays prime involvement in flavonoid biosynthesis, while ReCHS2 seems more flexible toward substrate selectivity and may be implicated in the biosynthesis of polyketidic anthraquinones.

RESULTS

Isolation of Complementary DNA Clones of ReCHSs

Degenerate primers and the RACE PCR strategy were used to get the complete coding sequences of ReCHS paralogs. The nucleic acid sequence alignment revealed a high level of sequence similarity to related plant CHSs using BLASTN/BLASTX analysis tools. The isolated genes designated as *ReCHS1* (GenBank accession no. KF850684) and *ReCHS2* (GenBank accession no.

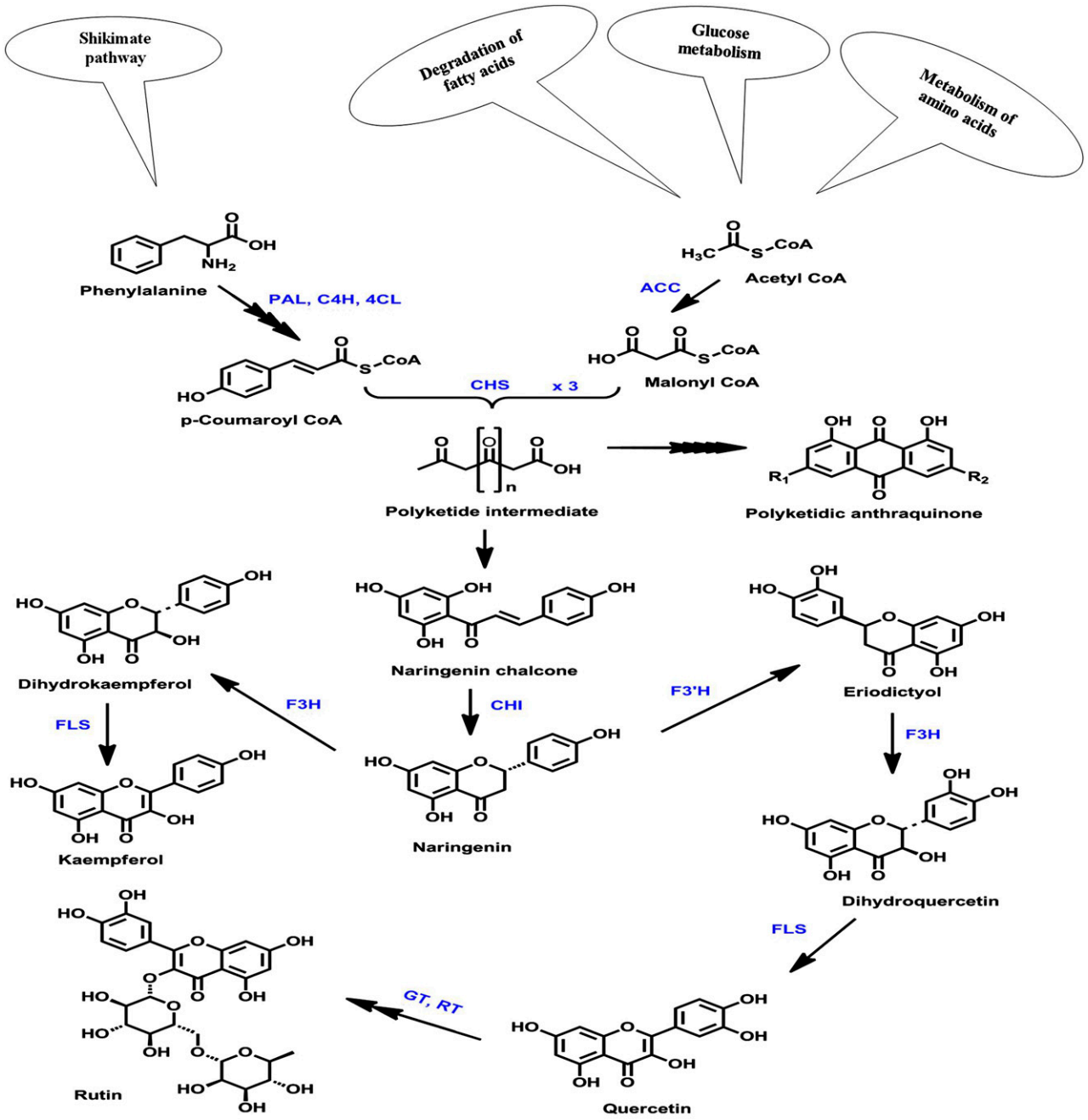

Figure 1. Flavonoid biosynthetic pathway. This schematic representation of part of the flavonoid biosynthetic pathway illustrates the biosynthesis of major flavonoid constituents. PAL, Phe ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumaroyl-CoA ligase; ACC, acetyl-CoA carboxylase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; FLS, flavonol synthase; GT, glucosyltransferase; RT, rhamnosyltransferase.

KC822472), with an open reading frame (ORF) of 1,179 bp each, also were found to be intronless. However, the main difference was observed in their untranslated regions, implicated to be important for the control of gene expression in plants at the posttranscriptional level. Moreover, the amino acid sequences encoded by full-length cDNAs of ReCHSs were shown to exhibit extended sequence similarity (75%–82%) with orthologous sequences of related polygonacious members, including *Rheum palmatum* (GenBank accession no. ABB13607.1), *Polygonum cuspidatum* (GenBank accession no. AFD64563.1), *Fallopia multiflora* (GenBank accession no. ADK45325.1), *Fagopyrum tataricum* (GenBank accession no. ACZ51475.1), and *Fagopyrum esculentum* (GenBank accession no. ADT63062.1), using the BLASTX/BLASTP algorithm.

In Silico Characterization and Phylogenetic Analyses

The ORFs of ReCHS1 and ReCHS2 were subjected to translation to generate 392 amino acids, corresponding to protein of 43.23 and 43.65 kD with calculated pI values of 6.03 and 9.21, respectively. The pairwise alignment of the deduced primary structures revealed that the two CHS homologs showed 83% identity at the nucleotide level and 75% at the amino acid level. The secondary structure analysis revealed that ReCHS1 and ReCHS2 are predominantly α-helical proteins with respective percentages for α-helices (37.76% and 41.58%), random coils (27.04% and 32.91%), β-turns (11.99% and 10.97%), and extended strands (19.39% and 18.37%). The ReCHS protein sequences were seen to lack any signal peptides and, thus, the transmembrane helices normally associated with the hydrophobic proteins. Analysis of the evolutionary conservation of ReCHS1 and ReCHS2 amino acid sequences revealed various high-score structural residues to be functional (Supplemental Fig. S1). Phyre2-based homology modeling was performed with a 100% confidence level, a coverage score of 98% for both ReCHSs, and percentage identity of 80% for ReCHS1 and 67% for ReCHS2 when 386 and 387 residues were aligned, respectively (Supplemental Fig. S2). The 3DLigandSite tool predicted 40 and 27 ligand-binding residues in ReCHS1 and ReCHS2, respectively (Supplemental Fig. S2, A and B). The I-TASSER-based model of ReCHS1 showed a confidence score of 1.72, 0.96 ± 0.05 template modeling score (estimated accuracy of model), and root mean square deviation (RMSD) of 3.3 ± 2.3 Å. Similarly, the confidence score, template modeling score, and RMSD for ReCHS2 were 1.71, 0.95 ± 0.05, and 3.3 ± 2.3 Å, respectively (Supplemental Fig. S2, C and D). The 3D models of both proteins when superimposed were found to exhibit a higher level of similarity ($P = 0.00$) with that of the template. The RMSD for ReCHS1 was 0.47 Å and that for ReCHS2 was 0.41 Å, with no twists found (Supplemental Fig. S2, E and F).

The well-known conservation of CHS sequences across species was exploited to identify the catalytically important residues in ReCHSs using the Clustal Omega tool. The conserved amino acid residues present in almost all type III PKSs also were found to be preserved in the primary amino acid sequences of ReCHSs. Multiple sequence alignment revealed that ReCHSs maintain the identical conserved catalytic triad Cys-164, His-303, and Asn-336 (numbering in alfalfa CHS) and a highly conserved Phe residue, Phe-216 (marked with yellow background; Fig. 2). ReCHSs also contain 13 inert active site residues (marked with cyan background) that shape the geometry of the active site, a malonyl-CoA binding motif, VET/AKLGLKEEKLKATRQ (marked with green background), and a highly conserved signature sequence, GVLFGF (marked with purple background). In addition, the seven amino acid residues (Thr-132, Met-138, Phe-216, Ile-255, Gly-257, Phe-266, and Pro-376) involved in the formation of the cyclization pocket also were found to be conserved in both the CHS sequences (not highlighted in Fig. 2). However, the amino acids Ile-255 and Gly-257 seem to be nonsynonymously replaced by Leu and Ala, respectively, in ReCHS2.

To elucidate the phylogenetic relationship of deduced primary amino acid sequences of ReCHSs with related CHS proteins, phylogenetic analysis was performed. About 30 amino acid sequences were selected after scrutinizing all the sequences at the order (Caryophyllales) level related to CHS genes submitted to the National Center for Biotechnology Information (NCBI) database. The sequences selected to ascertain the degree of evolutionary relatedness were chosen from 33 families (as per the APG III system, 2009) based solely on the complete coding sequence information available (GenBank). The phylogenetic tree was rooted using CHS from *Paeonia lactiflora* (Chinese peony) as an outgroup, considering that it belongs to the Saxifragales order (Paeoniaceae family), which shares a close evolutionary relationship with Caryophyllales. The results obtained showed that Caryophyllales CHS members exhibited a different evolutionary history. The CHS sequences from *Beta vulgaris* clustered with the homologous CHSs from the same species, signifying their late-diverged evolution. *Silene latifolia* also clustered with CHSs from *B. vulgaris*. Two species of *Dianthus* formed a separate clade. Moreover, the two members of ReCHS clustered with different orthologous members of the same family in two separate and distinct clades. Additionally, a similar pattern of selected CHS sequences was obtained in both the DNA-based and amino acid-based trees (Fig. 3). Additionally, the calculation of synonymous and nonsynonymous substitution rates in ReCHS nucleotide sequences showed that the number of synonymous mutations was much higher compared with that of nonsynonymous ones (Supplemental Table S1).

Heterologous Expression, Purification, and Bottom-Up Proteomic Analysis of Recombinant ReCHSs from *E. coli* Cellular Extracts

The identity of the isolated cDNA clones of ReCHSs was confirmed by the production of recombinant proteins in *E. coli* BL21 (DE3). The biochemical characterization of gene products was accomplished using an isopropylthio-β-galactoside (IPTG)-inducible *E. coli* expression vector, pGEX-4T-2, under the control of the Ptac hybrid promoter. The highest expression level for each of the generated constructs was observed at 0.8 mM IPTG induction for 6 h at 30°C. The recombinant enzymes displayed an apparent molecular mass of approximately 69 kD, which was similar to the predicted mass of recombinant fusion proteins, including glutathione *S*-transferase (GST; 25.99 kD). The optimum expression level was selected for further purification of respective proteins using the principle of affinity chromatography. The respective fusion proteins were reduced to their normal homodimeric

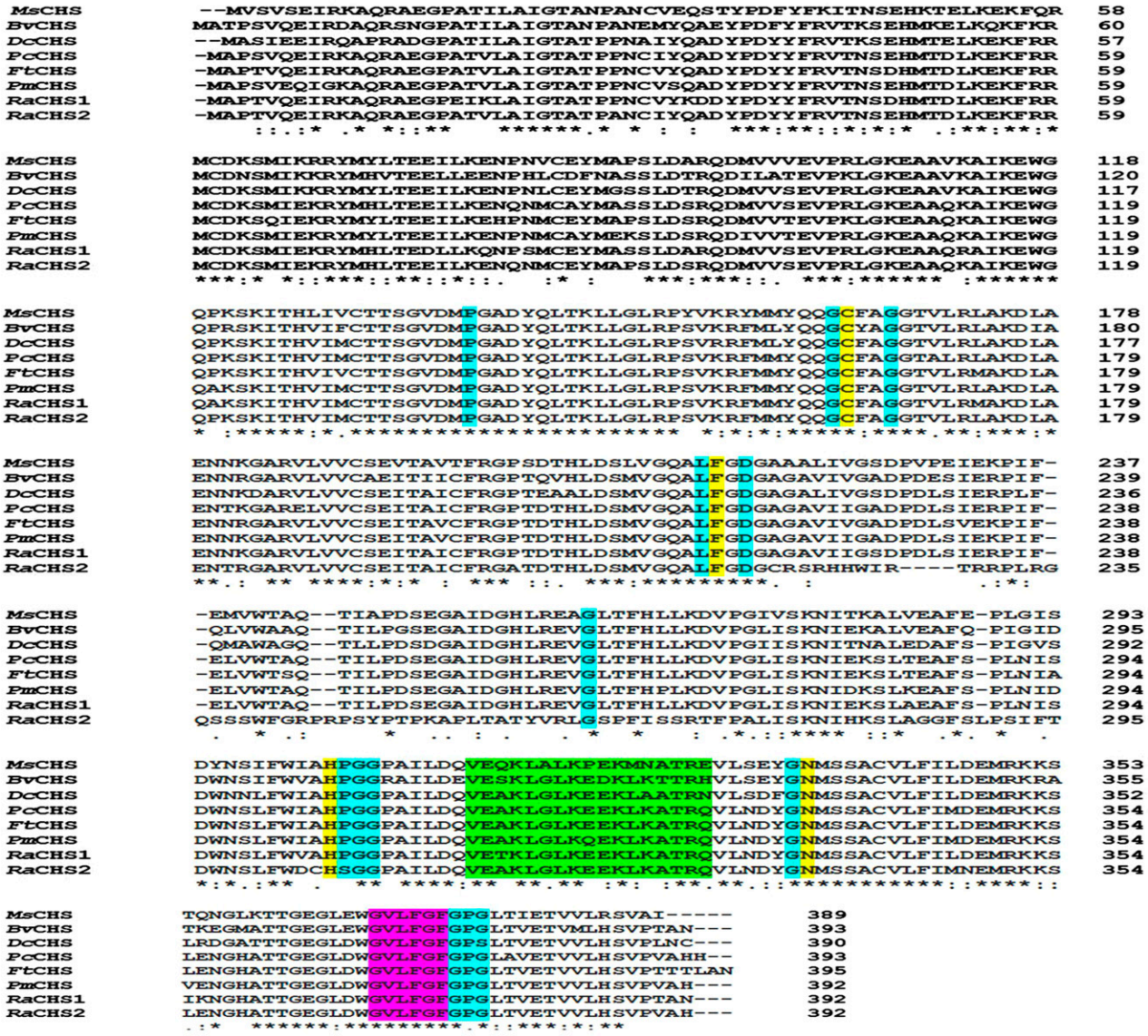

Figure 2. Multiple sequence alignment of the deduced amino acid sequences of ReCHS1 and ReCHS2 with related plant CHS sequences using the Clustal Omega multiple sequence alignment tool. Amino acid positions are given on the right, and identical, conserved, and semiconserved amino acids are indicated with asterisks, colons, and periods, respectively. Functionally important conserved residues are highlighted with a colored background: yellow, the four catalytic residues (Cys-His-Asp triad + Phe) that are shown to be conserved in all polyketide synthases; cyan, the 13 residues that shape the geometry of the active site; green, the malonyl-CoA-binding motif; and purple, the highly conserved CHS signature sequence, *N*-myristoylation motif. Abbreviations (with GenBank accession numbers) are as follows: MsCHS, *Medicago sativa* CHS (L02902); BvCHS, *Beta vulgaris* CHS (XM_010693892.1); DcCHS, *Dianthus caryophyllus* CHS (Z67982.1); PcCHS, *Polygonum cuspidatum* CHS (EF090266.2); FtCHS, *Fagopyrum tataricum* CHS (HQ434624.1); PmCHS, *Persicaria minor* CHS (JQ801338.1); RaCHS1, *Rheum australe* CHS1 (KF850684); and RaCHS2, *Rheum australe* CHS2 (KC822472).

size (approximately 42 kD) by cleavage at the thrombin site located toward the C terminus of the GST tag (Fig. 4).

The bottom-up mass spectrometry approach of proteomic analysis is the most popular method in large-scale proteomics. The sequence coverage of 63.01% for ReCHS1 and 48.21% for ReCHS2 was obtained using Agilent Mass Hunter BioConfirm version B.06.00 software. The tryptic digests of purified ReCHSs provided accurate masses as well as tandem mass spectra of different peptides with reasonable mass accuracy. The tandem mass spectrometry (MS/MS) characterization of molecular ion species revealed the identification of five and six tryptic peptides, respectively, for ReCHS1 and ReCHS2, with the length of peptides varying from four to 23 amino acids (Table I). Fragmentation

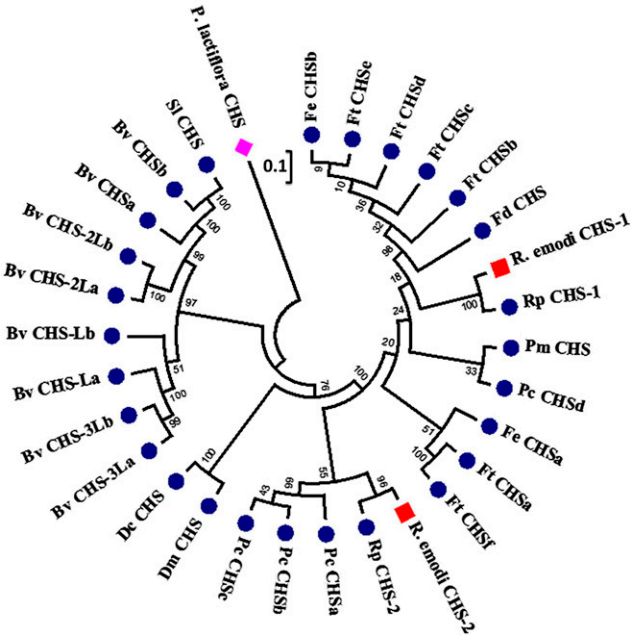

Figure 3. Phylogenetic tree of ReCHS1 and ReCHS2. The phylogenetic analysis was performed using the MUSCLE program and MEGA6 software based on the neighbor-joining method. The numbers on the nodes indicate the bootstrap values after 1,000 replicates. The bar indicates an evolutionary distance of 0.01%. The evolutionary distances were computed using the Poisson correction method. The analysis involved the alignment of 30 amino acid sequences that were chosen by scrutinizing the available data related to CHS genes from the NCBI database at the order level (Caryophyllales). About 33 families (as per the APG III system, 2009) were screened, and the desired sequences were selected based on the complete coding sequence information available. The phylogenetic tree was rooted using CHS from Chinese peony (*Paeonia lactiflora*; AEK70334.1) as an outgroup, seeing that it belongs to the Saxifragales order (Paeoniaceae family), which shares a close evolutionary relation with Caryophyllales. The numbers on the branches indicate the bootstrap values after 1,000 replicates. The database accession numbers of the CHS sequences used are as follows: *Fagopyrum tataricum* (FtCHSa, AIY62394.1; FtCHSb, AHU87068.1; FtCHSc, ADU05554.1; FtCHSd, ADL39795.1; FtCHSe, ACZ51475.1; FtCHSf, ACH70135.1); *Fagopyrum dibotrys* (FdCHS, ACZ48699.1); *Fagopyrum esculentum* (FeCHSa, ADT63062.1; FeCHSb, ACZ51476.1); *Rheum palmatum* (RpCHS-1, ABB13607.1; RpCHS-2, ABB13608.1); *Rheum australe* (*R. australe* CHS-1, AHC28523.1; *R. australe* CHS-2, AHB19194.1); *Persicaria minor* (PmCHS, AFI98395.1); *Polygonum cuspidatum* (PcCHSa, AFD64563.1; PcCHSb, ABK92282.2; PcCHSc, ABK92281.2; PcCHSd, ACC76754.1); *Dianthus monspessulanus* (DmCHS, AAF81743.1); *Dianthus caryophyllus* (DcCHS, CAA91923.1); *Beta vulgaris* ssp. *vulgaris* (BvCHSa, XP_010692112.1; BvCHSb, XP_010690460.1; BvCHSLa, XP_010692195.1; BvCHSLb, XP_010683006.1; BvCHS-2La, XP_010670607.1; BvCHS-2Lb, XP_010670606.1; BvCHS-3La, XP_010692194.1; BvCHS-3Lb, XP_010692193.1); and *Silene latifolia* (SlCHS, BAE80096.1).

during mass spectrometry analysis yielded numerous b- and y-type ions, which permitted sequencing of the peptide. In general, the quadrupole time-of-flight mass spectrometry spectra of the recognized peptides gave an error of less than 1 ppm from that of the observed theoretical mass (m/z ratio). The detailed information regarding the extracted EIC spectra, their molecular ion peaks, and the generated collision-induced dissociation MS/MS fragmentation patterns with y- and b-type fragments (according to the nomenclature of Domon and Costello [1988]) of the corresponding identified peptides are depicted in Supplemental Figures S3 and S4.

Functional Characterization and Enzymatic Properties of ReCHSs

To investigate the kinetic properties of ReCHSs, the known concentrations of purified proteins were tested with *p*-coumaroyl-CoA (a common starter unit) and malonyl-CoA (an extender unit) as substrates, and the reaction products were analyzed by liquid chromatography-mass spectrometry (LC-MS) in comparison with authentic standards of naringenin and naringenin chalcone. Incubation of ReCHSs with *p*-coumaroyl-CoA and malonyl-CoA generated the product(s) with expected retention times of 14.9 and 13.8 min, respectively, for naringenin and naringenin chalcone, confirming that the isolated cDNAs encoded CHS with typical enzymatic function (Fig. 5, A–C). The positive electrospray ionization (ESI)-ion mass spectrum resolved a molecular ion $[M-H]^+$ at m/z of 273, similar to that of the reference compounds as depicted in the MRM graphs generated (Fig. 5). Moreover, the profile of the fragmented form of $[M-H]^+$ with the MRM transition masses of m/z 273/153 and m/z

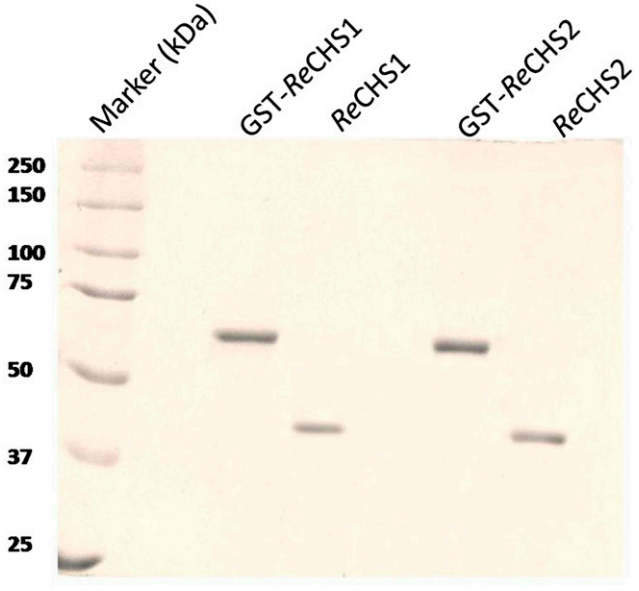

Figure 4. SDS-PAGE profile of purified recombinant proteins. SDS-PAGE (10%) is shown for affinity-purified recombinant proteins from *E. coli* BL21 (DE3) cells transformed with pGEX-ReCHS1 and pGEX-ReCHS2 expression cassettes. Lane 1, Standard protein marker; lane 2, purified recombinant GST-fused ReCHS1 protein; lane 3, purified ReCHS1 protein after removal of GST; lane 4, purified recombinant GST-fused ReCHS2 protein; lane 5, purified ReCHS2 protein after removal of GST.

Table I. *List of fragment ions generated in the mass spectrum and tandem mass spectrum related to the peptide identification from the tryptic digests of ReCHS1 and ReCHS2 proteins*

Protein	Retention Time	Precursor Ion	Charge (Ionization Mode)	Theoretical Mass	Experimental Mass	Error	Sequence Location	Peptide Sequence
	min			*m/z*	*m/z*	*ppm*		
ReCHS1	12.42	630.7983 $[M+2H]^{+2}$	2 (+)	1,259.5822	1,259.5816	0.48	A44–A54	VTNSDHMTDLK
	15.46	580.2929 $[M+2H]^{+2}$	2 (+)	1,158.5708	1,158.5703	0.4	A96–A105	QDMVVSEVPR
	16.33	828.4812 $[M+H]^{+1}$	1 (+)	827.4737	827.4753	−1.94	A271–A278	DVPGLISK
	17.38	847.4048 $[M+3H]^{+3}$ 1,270.600 $[M+2H]^{+2}$	3 (+)	2,537.1835	2,537.1964	−5.08	A125–A147	ITHVIMCTTSGVDMPGADYQLTK
	18.24	421.5535 $[M+3H]^{+3}$ 631.8258 $[M+2H]^{+2}$	3,2 (+)	1,261.6366	1,261.6377	−0.8	A70–A79	YMHLTEDLLK
ReCHS2	5.61	553.2109 $[M+H]^{+1}$ 575.1909 $[M+Na]^{+1}$	1,1 (+)	552.2030	552.2036	−1.09	A60–A63	MCDK
	13.12	430.3021 $[M+H]^{+1}$	1 (+)	429.2944	429.2951	−1.56	A318–A321	LGLK
	13.51	607.309 $[M+H]^{+1}$	1 (+)	606.3044	606.3047	−1.94	A64–A68	SMIEK
	15.46	491.8245 $[M+2H]^{+2}$ 982.6379 $[M+H]^{+1}$	2,1 (+)	981.6338	981.6335	0.31	A148–A156	LLGLRPSVK
	18.55	482.2695 $[M+2H]^{+2}$ 963.5290 $[M+H]^{+1}$	2,1 (+)	962.5235	962.5185	5.13	A262–A270	LGSPFISSR
	18.96	933.9299 $[M+2H]^{+2}$	2 (+)	1,864.8413	1,864.8423	−0.5	A158–A173	FMMYQQGCFAGGTVLR

273/147 was almost similar to that of the standards (Fig. 5, D and E). It was also observed that the reactions quenched by acidification showed the formation of naringenin only, while the nonacidified ones generated naringenin as well as naringenin chalcone, so two cyclic isoforms with the same molecular mass had different retention times (Fig. 5F).

The type III PKSs are commonly known to possess broad substrate tolerance; therefore, we investigated the starter unit specificity of two ReCHS paralogs. Detailed kinetic studies were carried out using five different starter units, the results of which are summarized in Table II. The V_{max} values for *p*-coumaroyl-CoA starter, as calculated by nonlinear regression analysis, were 42 and 37 pmol min^{-1} mg^{-1}, whereas the apparent K_m values were 38.43 and 155.8 μM, respectively, for ReCHS1 and ReCHS2 (Fig. 6, A and B; Table II). In general, K_m was seen to improve when using other CoA esters, but significant changes were observed in V_{max} values of both enzymes. ReCHS1 and ReCHS2 displayed poor substrate affinity toward hexanoyl-CoA and *p*-coumaroyl-CoA, respectively, as reflected in their K_m values (Table II). ReCHS1 also exhibited higher efficiency (V_{max}/K_m) with *p*-coumaroyl-CoA, and the V_{max}/K_m values for the rest of the starter units were nearly the same (Table II). However, V_{max}/K_m values showed a significant increase in the case of ReCHS2, where its overall range among all the starter substrates except *p*-coumaroyl-CoA was much higher than that of ReCHS1 (Table II). Furthermore, the relative activities with different starter CoA esters were plotted for each enzyme, as depicted in Figure 7. The relative activity of ReCHS paralogs yielded interesting results, with ReCHS2 exhibiting higher activity with nonphysiological substrates compared with ReCHS1.

Additionally, ReCHS1 was found to be active over a wide range of pH values used, whereas ReCHS2 showed a very short range of activity, as depicted in Figure 6C. Monitoring the catalytic activity of enzymes with different buffer systems over a gradient of pH values (pH 5–9) revealed that the optimal pH for ReCHS1 ranged from 7 to 7.5, while for ReCHS2, the optimal pH was 7. Meanwhile, altering pH had little effect on the observed activity of ReCHS1 compared with that of ReCHS2, which is rendered inactive over a set of pH values (Fig. 6C).

Genomic Southern-Blot Analysis

A genomic Southern-blot analysis was performed in order to estimate the number of ReCHS gene copies and further validate the two isoforms. Each enzyme nearly produced a restriction pattern that was consistent with a single-copy gene. For ReCHS1, a single band was observed for DNA digested with *Bam*HI and two when digested with *Nco*I. Similarly, ReCHS2 scored one band with *Eco*RI and two with *Xma*I (Fig. 8). The results obtained suggest that the genome of *R. emodi* contains two CHS paralogs with a single copy for each of them.

Expression Pattern of ReCHS Genes and Flavonoid Accumulation

To understand the spatial regulation of the ReCHS genes in *R. emodi*, the expression pattern of ReCHS1 and ReCHS2 in different tissues was examined using relative quantitative real-time PCR. The transcripts of ReCHS genes were detected in all the examined samples with a distinct expression pattern. ReCHS1 transcript levels were higher in root, followed by leaf and stem, whereas ReCHS2 transcripts were more obvious in stem, followed by leaves and root (Fig. 9A). The differential transcript levels of ReCHSs in different organs were in agreement with earlier studies of the CHS multigene families of *G. max* and *G. hybrida* (Tuteja

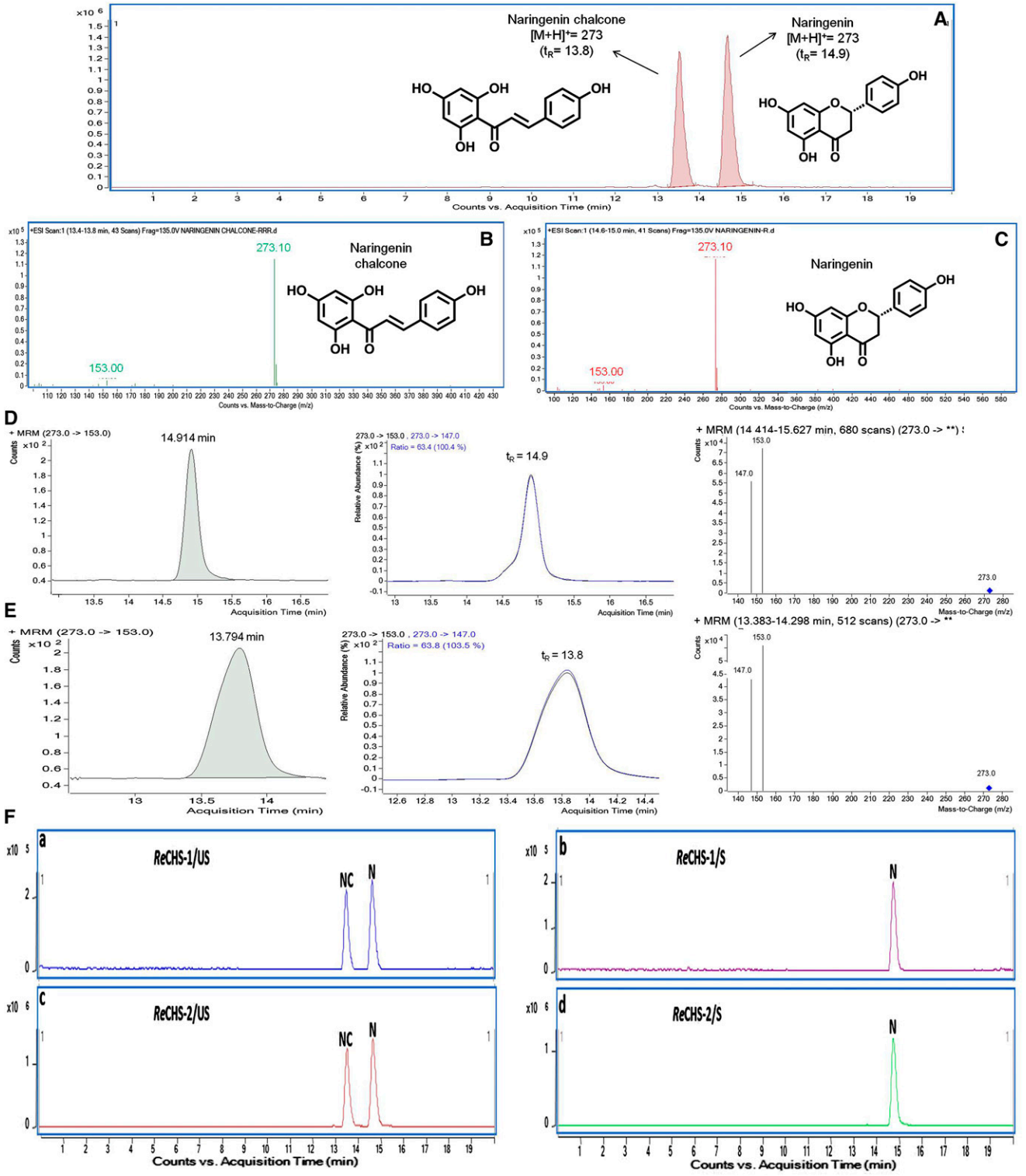

Figure 5. Multiple reaction monitoring (MRM) graphs. A, MRM chromatograms of the standard compounds naringenin and naringenin chalcone eluting at 14.9 and 13.8 min, respectively. B and C, Mass spectrometry spectra of naringenin chalcone (B) and naringenin (C). D, MRM graphs of naringenin with transition masses of *m*/*z* 273/153 and 273/147. E, MRM graphs of naringenin chalcone with transition masses of *m*/*z* 273/153 and 273/147. F, Liquid chromatography-MRM graphs of unstopped and stopped in vitro enzyme reactions of ReCHS1 (a and b) and ReCHS2 (c and d) enzymes. N, Naringenin; NC, naringenin chalcone; S, stopped; US, unstopped.

Table II. *Steady-state kinetic parameters of purified ReCHSs with different starter units*

Results are means (n = 3) with SD values below 10% in all the cases.

Starter CoA	ReCHS1			ReCHS2		
	K_m	V_{max}	Efficiency (V_{max}/K_m)	K_m	V_{max}	Efficiency (V_{max}/K_m)
	μM	*pmol min^{-1} mg^{-1}*		μM	*pmol min^{-1} mg^{-1}*	
p-Coumaroyl-CoA	38.43	42	1.092	155.8	37.0	0.237
Acetyl-CoA	46.6	3.23	0.069	37.3	63.5	1.702
Butyryl-CoA	13.9	1.0	0.071	8.01	11.0	1.373
Hexanoyl-CoA	91.65	1.27	0.013	11.4	4.7	0.412
Octanoyl CoA	22.0	1.1	0.050	26.6	21.4	0.804

et al., 2004; Deng et al., 2014). In general, the expression pattern of ReCHSs was coincident with flavonoid accumulation. The content of naringenin was found to be higher in roots (4.671 ± 0.439 mg g^{-1}), whereas leaves showed the maximum accumulation of rutin (12.896 ± 0.989 mg g^{-1}; Fig. 9G; Supplemental Table S2). A similar trend for rutin accumulation was reported in *F. esculentum*, which has been well studied for rutin and its biosynthetic machinery (Li et al., 2010).

To ascertain the role of ReCHSs in secondary metabolite biosynthesis in *R. emodi*, the expression levels of the two paralogs also were investigated in plant samples collected from four different geographic locations spread over an altitudinal range of 1,600 to 4,500 m above sea level (asl). The transcript levels of two CHS genes differ widely in different altitudes, with both genes showing highest expression in the Nyoma Valley, Ladakh (33° 08′ 661′′ N, 78° 34′ 742′′ E; 4,415 m asl; Fig. 9B). Flavonoids exhibited a general increasing trend with increasing altitude, particularly two predominant flavonoids, naringenin and rutin, accumulating at higher concentrations. The concentration of naringenin ranged from 7.431 ± 0.443 to 15.266 ± 1.043 mg g^{-1} and that of rutin from 11.603 ± 1.413 to 17.887 ± 1.114 mg g^{-1} (Fig. 9H; Supplemental Table S2). Our observations were in conformity with earlier reports on *Arnica montana* (Spitaler et al., 2006) and *F. tataricum* (Guo et al., 2011).

In addition, naringenin and rutin (quercetin-3-*O*-rutinoside) were found to be the major flavonoid compounds in all the samples analyzed, whereas kaempferol and quercetin were found in the least abundance, with the former ranging from 0.0068 ± 0.001 to 0.1637 ± 0.022 mg g^{-1} and the latter from 0.0309 ± 0.004 to 3.915 ± 0.879 mg g^{-1} (Fig. 9; Supplemental Table S2). Like many plant natural products, the shift of quercetin toward its glycosylated form rutin may be aimed at increasing their solubility and stability for their easy storage and accumulation in plant cells (Farooq et al., 2013).

Isolation and in Silico Characterization of ReCHS Promoter Sequences

The expression of CHS genes can be induced by various biotic and abiotic elicitations, including light/UV light, phytopathogens, mechanical wounding, and plant hormones (Dao et al., 2011). These modulate gene expression by interacting with the cis-regulatory elements in the promoter regions. To elucidate the transcriptional regulation of two ReCHS paralogs, the respective 5′ flanking regions were identified and further examined in silico for various putative cis-regulatory promoter elements. The genome-walking approach led to the identification of 413- and 388-bp promoter regions of ReCHS1 and ReCHS2, respectively. The predicted transcription initiation site (+1) was found to be located at 62 and 75 bp upstream of the start codon, whereas the putative TATA box was 41 and 45 bp upstream of the transcription initiation site in ReCHS1 and ReCHS2, respectively (Supplemental Fig. S5, A and B). In silico analysis of the isolated promoter sequences was carried out by the PLACE and PlantCare databases. Several important cis-acting regulatory elements were identified within the promoter regions of ReCHSs (Table III). Four regulatory elements (G box, CGTCA motif, TGACG motif/TCA element, and W box) were chosen to investigate their role in stress responsiveness, with the aim to study the inducible/repressible nature of the regulatory motifs.

Effects of Abiotic Elicitors on ReCHS Expression vis-à-vis Flavonoid Biosynthesis

Elicitors selected on the basis of promoter analysis were evaluated with regard to ReCHS expression pattern vis-à-vis flavonoid and anthraquinone accumulation. The experiments were conducted to compare the endogenous response induced via wounding and the exogenous induction via the application of MeJ (0.1 mM), SA (0.1 mM), and UV-B light (1,500 μJ m^{-2}) exposure.

Jasmonic acid and its esters like MeJ have long been reported to play a signaling role in insect and disease resistance. MeJ treatment significantly induced the transcript levels of ReCHS1 and ReCHS2, reaching the highest levels at 12 and 24 h with nearly 15- and 12-fold increases, respectively (Fig. 9C). Afterward, the transcript levels of both the paralogs declined. These results are in conformity with earlier studies on *Picea glauca* (Richard et al., 2000) and *Plagiochasma appendiculatum* (Yu et al., 2015). Earlier studies on *Rubus* spp. demonstrated that MeJ significantly enhances total flavonoid content in blackberry fruit and also displayed a positive correlation with the increasing concentrations of MeJ used for the treatment (Wang et al., 2008). We also

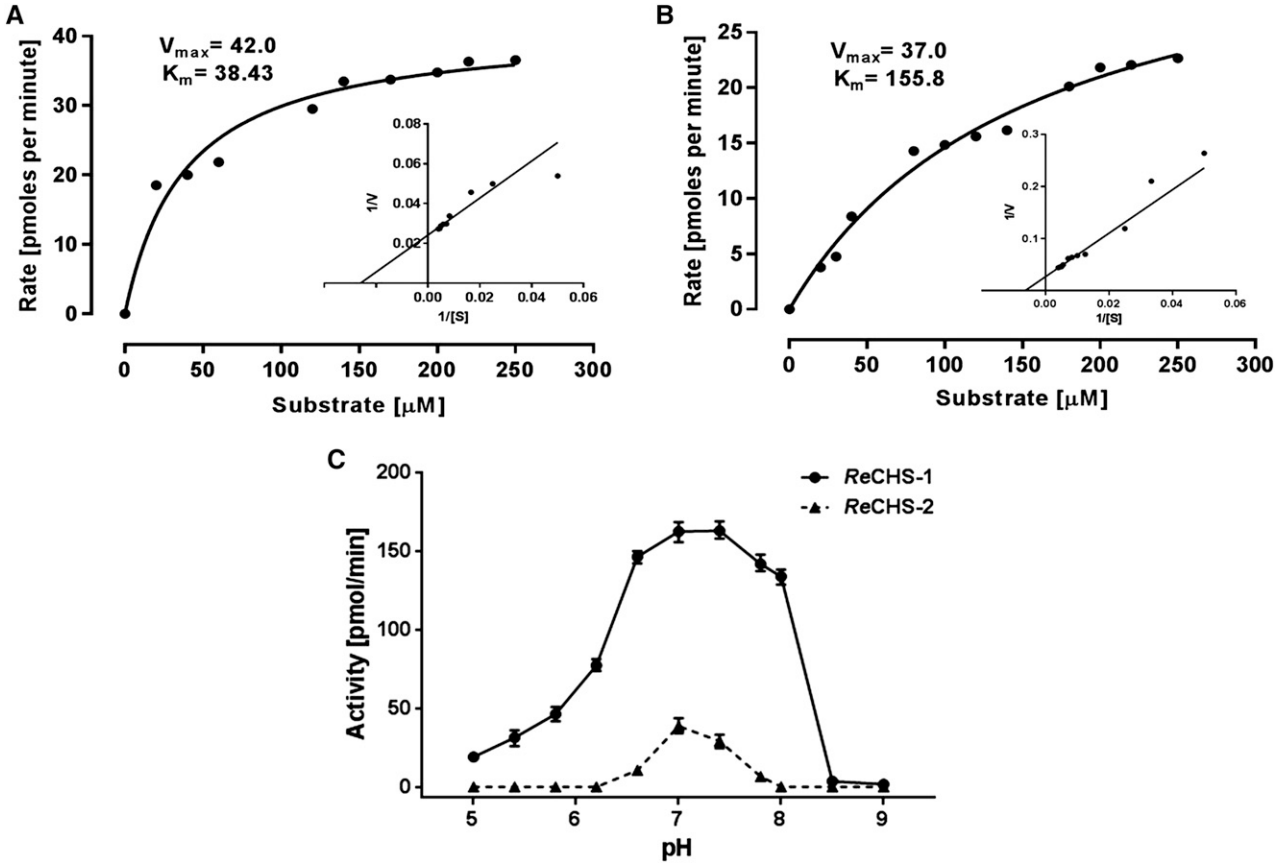

Figure 6. Kinetic study of ReCHSs. A and B, Michaelis-Menten plots of ReCHS1 (A) and ReCHS2 (B) with inset Lineweaver-Burk plots. The kinetic parameters K_m and V_{max} were calculated by nonlinear regression analysis using GraphPad Prism 6 software. C, The activity of ReCHSs was assayed at varied pH values (pH 5–9). Citrate buffer, potassium phosphate buffer, and 0.1 M Tris-HCl buffers were used for pH 3 to 6.2, 5.8 to 8, and 8.5 to 9, respectively. The CoA esters (*p*-coumaroyl-CoA and malonyl-CoA) were used as substrates, and the production of naringenin/naringenin chalcone was quantified as activity (pmol min^{-1}). Values are the means ± SD of at least three replicates. Points of variance (n = 3) are depicted at each point in C.

observed a similar effect of MeJ on elicitor-treated tissues of *R. emodi*, wherein the major flavonoids, naringenin and rutin, accumulated to concentrations of 8.507 ± 0.637 mg g^{-1} (6-fold increase compared with that of control) and 17.086 ± 0.321 mg g^{-1} (1.5-fold increase), respectively, at 24 h. The concentrations of both flavonoids dropped at 48 h (Fig. 9I; Supplemental Table S2).

The widely distributed phytohormone SA plays an important role in plant defense reactions against a broad range of stresses through morphological, physiological, and biochemical mechanisms. Here, SA elicited differential increases in the expression levels of ReCHS1 and ReCHS2. The respective paralogs registered 3-fold enhanced transcript levels in 12-h and 12- to 24-h induction periods (Fig. 9D). Two predominant flavonoids exhibited a similar trend wherein naringenin and rutin showed higher accumulation at 24 h. The respective contents of naringenin and rutin were nearly 2.5-fold (3.6474 ± 0.451 mg g^{-1}) and 1.5-fold (16.402 ± 0.817 mg g^{-1}) higher than that of the control (Fig. 9J; Supplemental Table S2). SA treatments also have resulted in significant increases in the flavonoid content in root suspension cultures of *Panax ginseng* (Ali et al., 2007).

As UV-B (280–320 nm) radiation is an important abiotic environmental factor inducing flavonoid biosynthesis, the in vitro-raised plantlets were subjected to UV light treatment to study the expression levels of ReCHSs. It was first shown in Arabidopsis (*Arabidopsis thaliana*) that UV and blue light could induce the expression of CHS genes (Jenkins et al., 2001). The plant tissues treated with UV light showed a gradual increase in the expression of ReCHS1 up to a maximum of 2-fold at 9 h post induction, whereas a steep increase was observed in the transcript level of ReCHS2, which showed a nearly 5-fold increase at 6 h followed by a 12-fold increase at 9 h (Fig. 9E). These results were in conformity with earlier studies in which a 10-fold increase in CHS expression of mature Arabidopsis leaves was reported (Jenkins et al., 2001). The lower transcript level of ReCHS1 as compared with that of ReCHS2 correlates with earlier studies on *G. max* in which only the CHS1 member of the gene family showed induction

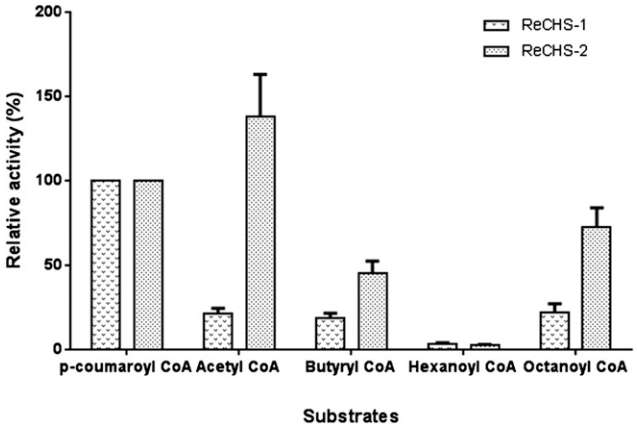

Figure 7. Relative activity. The relative activities of purified ReCHSs were observed using 120 μM malonyl-CoA and 200 μM starter-CoA esters. The activity percentage was based on the production of naringenin. The activity of either enzyme with *p*-coumaroyl-CoA was taken as 100%, and respective relative activities with different starter units are plotted for each enzyme. Values are means ± SD of at least three replicates.

upon UV irradiation (Tuteja et al., 2004). Flavonoids are reported to be UV-B light-absorbing compounds that show increased accumulation in plant cells when exposed to UV-B irradiation (Tossi et al., 2012). A parallel increasing trend was observed in UV-B light-elicited tissues of *R. emodi*. Both naringenin (5.0182 ± 0.66 mg g^{-1}) and rutin (16.705 ± 0.732 mg g^{-1}) were found to exhibit nearly 2-fold increases compared with the control at 9 and 6 h, respectively (Fig. 9K; Supplemental Table S2).

The endogenous responses of plants were observed via multiple wounding, and the effects were found to be similar to that of MeJ. Nevertheless, a stronger effect was observed, with a significantly different induction profile exhibiting a 4- to 5-fold increase in transcript levels of both ReCHSs at 24 h (Fig. 9F). It is generally thought that volatile jasmonates released from wounded tissues activate CHS genes, causing an advanced production of phytoalexins to resist any infection to the plant (Dao et al., 2011). A corresponding increasing pattern of flavonoids was observed in which both the major flavonoid constituents showed continuous increases in their accumulation until 48 h of incubation. Naringenin accumulated to a concentration of 5.3313 ± 0.883 mg g^{-1} and rutin to 17.569 ± 1.619 mg g^{-1}, displaying nearly 2-fold increases compared with the control (Fig. 9L; Supplemental Table S2). Enhanced flavonoid levels also have been reported in tissues subjected to UV-B light elicitation or mechanical wounding. In the GH variety of *Prunus persica*, wounding was found to augment the levels of flavonol and flavonoid contents (Tosetti et al., 2014).

To supplement the real-time and HPLC assays for respective gene expression and metabolite accumulation studies, we analyzed the changes in flavonol accumulation in situ in different plant tissues and elicitor-treated samples of *R. emodi* by DPBA staining. The fluorescent probe DPBA was used, and flavonoids were detected by microscopy. The abundance of flavonoids in photographed plant samples was correlated with the intensity of fluorescence generated by flavonol-conjugated DPBA. In a broader sense, flavonoid accumulation in terms of fluorescence was, by and large, in conformity with the quantitative HPLC data. Representative images of all the analyzed samples showing the in planta distribution of flavonoid accumulation by epifluorescence microscopy are depicted in Supplemental Figure S6.

R. emodi is also well recognized for being a rich repository of phytoconstituents called anthraquinones, which are well known for various biological activities. Against this backdrop, we extended our study to the determination of two major anthraquinone constituents: emodin and chrysophanol. In general, the concentration of chrysophanol was higher compared with emodin. The concentration of chrysophanol was shown to be elevated in all tissues, with maximum accumulation in leaf (3.707 ± 0.188 mg g^{-1}), followed by stem (3.338 ± 0.135 mg g^{-1}) and root (3.0198 ± 0.137 mg g^{-1}). Moreover, a general increasing trend of anthraquinones, especially chrysophanol, was observed in the elicitor-treated tissues, with increases in incubation time spread over a period of 12 to 48 h (3–9 h in the case of UV-B light-treated samples), as depicted in Figure 10 and Supplemental Table S2.

Correlation between ReCHS Transcript Levels and Metabolite Accumulation

The biological significance of metabolite-transcript correlations has usually been considered to reflect simple or complex associations between them (Mounet et al., 2009). We investigated the correlation between ReCHS gene expression profiles and the accumulation of metabolites (total flavonoids and total anthraquinones) to decipher the possible involvement of two divergent ReCHS paralogs toward either of the secondary metabolite groups. A positive significant correlation was found between ReCHS1 and ReCHS2 transcript levels and total flavonoid and total anthraquinone accumulation, respectively, vis-à-vis elicitor treatments. The tissue- and location-specific plant samples also exhibited a positive and nearly significant correlation between transcript profiles and metabolite content. Predominantly, correlation analysis between ReCHS1 transcript levels and anthraquinone accumulation showed a negative relation. Pearson's correlation analysis data are presented in Supplemental Tables S4 to S9.

DISCUSSION

Phenylpropanoid metabolism has been explored for the past two decades for possible biotechnological interventions due to its potential in agricultural and pharmaceutical applications (Wu and Chappell, 2008).

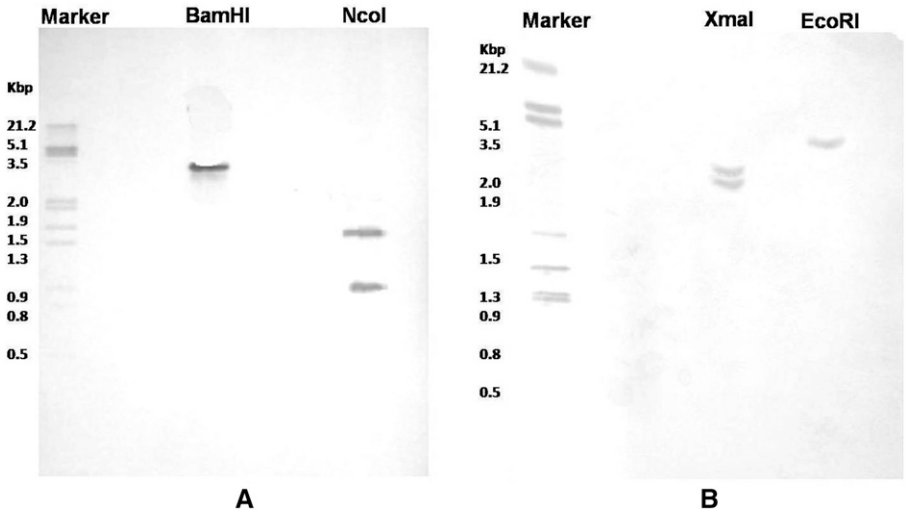

Figure 8. Southern-blot analysis of ReCHSs. The genomic DNA (greater than 20 μg) isolated from *R. emodi* was digested with *Bam*HI (noncutter) and *Nco*I (single cutter) for ReCHS1 (A) and *Eco*RI (noncutter) and *Xma*I (single cutter) for ReCHS2 (B). The digested samples were separated on 0.8% agarose gels, blotted onto a nylon membrane, and hybridized with digoxigenin-labeled ORFs of ReCHS1 and ReCHS2 as probes.

The ubiquitous occurrence, complex diversity, and diverse functions of flavonoids have made them suitable and effective targets for genetic engineering to alleviate the demands for limited natural resources. The in situ levels of major subgroups of flavonoids like chalcones and stilbenes have been increased because of their putative health benefits and prospective roles as defense agents of plants against various pathogens (Dixon et al., 1996). Owing to the endemic and endangered (Rokaya et al., 2012) nature of *R. emodi* and being a high-altitude medicinal herb not amenable to cultivation at low altitudes, it is quite indispensable to embark on a metabolic engineering program for enhanced and purposeful production of its characteristic phytoconstituents. With this viewpoint and as a prerequisite for heterologous and/or homologous production of flavonoids, two evolutionarily discrete paralogs belonging to the CHS superfamily of type III PKSs were isolated, fully characterized, and functionally validated from *R. emodi*. Our observations from real-time expression profiling, phytochemical evaluation, and kinetic studies were indicative of their functional promiscuity vis-à-vis flavonoid biosynthesis and substrate selectivity.

The highly conserved nature of CHS sequences across species was used to recognize catalytically important residues in ReCHS paralogs that showed modest similarity at the nucleotide and amino acid levels with each other and with related orthologous family members. Sequence analysis and homology modeling of the isolated ReCHSs revealed that they share similar attributes to those found in other known CHSs, like the representative alfalfa CHS2 (Jez et al., 2000). Additionally, the other identified residues (Fig. 2) thought to be essential in controlling the substrate and product specificity also were found to show significant levels of conservation in the isolated ReCHS sequences, which suggests that the two paralogs are true CHSs. Furthermore, the absence of signal peptides and transmembrane helices in both paralogs confirms their cytoplasmic localization, which is an established location for flavonoid biosynthesis. It seems plausible that ReCHS enzymes carry out the biosynthesis of naringenin chalcone directly in cytoplasm, which then isomerizes in vivo to naringenin.

Plants have evolutionarily employed CHS family members to cope with the shifting environment. It has been reported that the diversity of molecular evolutionary patterns of early- and late-diverged CHS members in different lineages are not related to enzyme function (Han et al., 2014). However, in our context, ReCHS1 showed a propensity for flavonoid biosynthesis, while ReCHS2 presented flexibility in terms of substrate selectivity vis-à-vis enzyme efficiency. It may be implicated in the biosynthesis of polyketidic anthraquinones. It is mainly inferred from the higher transcript levels of the ReCHS2 paralog in the elicitor-treated tissues correlating positively with the increased accumulation of anthraquinones.

The joint clustering of CHS sequences from *B. vulgaris* (Amarathaceae) suggests the possible late divergence of these homologs from their common ancestor. *S. latifolia* (Caryophyllaceae) was seen to cluster with the lineage of *B. vulgaris* CHS members, indicative of the probable evolutionary link between the two families. The appearance of ReCHS1 and ReCHS2 in two distinct and distant clusters points toward their early-diverged evolution. Moreover, both ReCHS1 and ReCHS2 members of *R. emodi* join their respective members from *R. palmatum* in a little clade in two distant branches. This is indicative of the two CHS variants (CHS1 and CHS2) having evolved before the speciation event between the two *Rheum* spp. In other words, CHS1 and CHS2 possibly diverged in some common ancestor of these two *Rheum* spp. The paralogs are distantly related and have evidently diverged in the ancestral lineage of the big clade that they belong to, as depicted in Figure 3. A similar evolutionary trend has been observed in *Sorbus*, *Phalaenopsis*, *Hypericum*, and *Bromheadia*, where CHS homologs diverged before the formation of these species (Han et al., 2014). Additionally, a recent study has

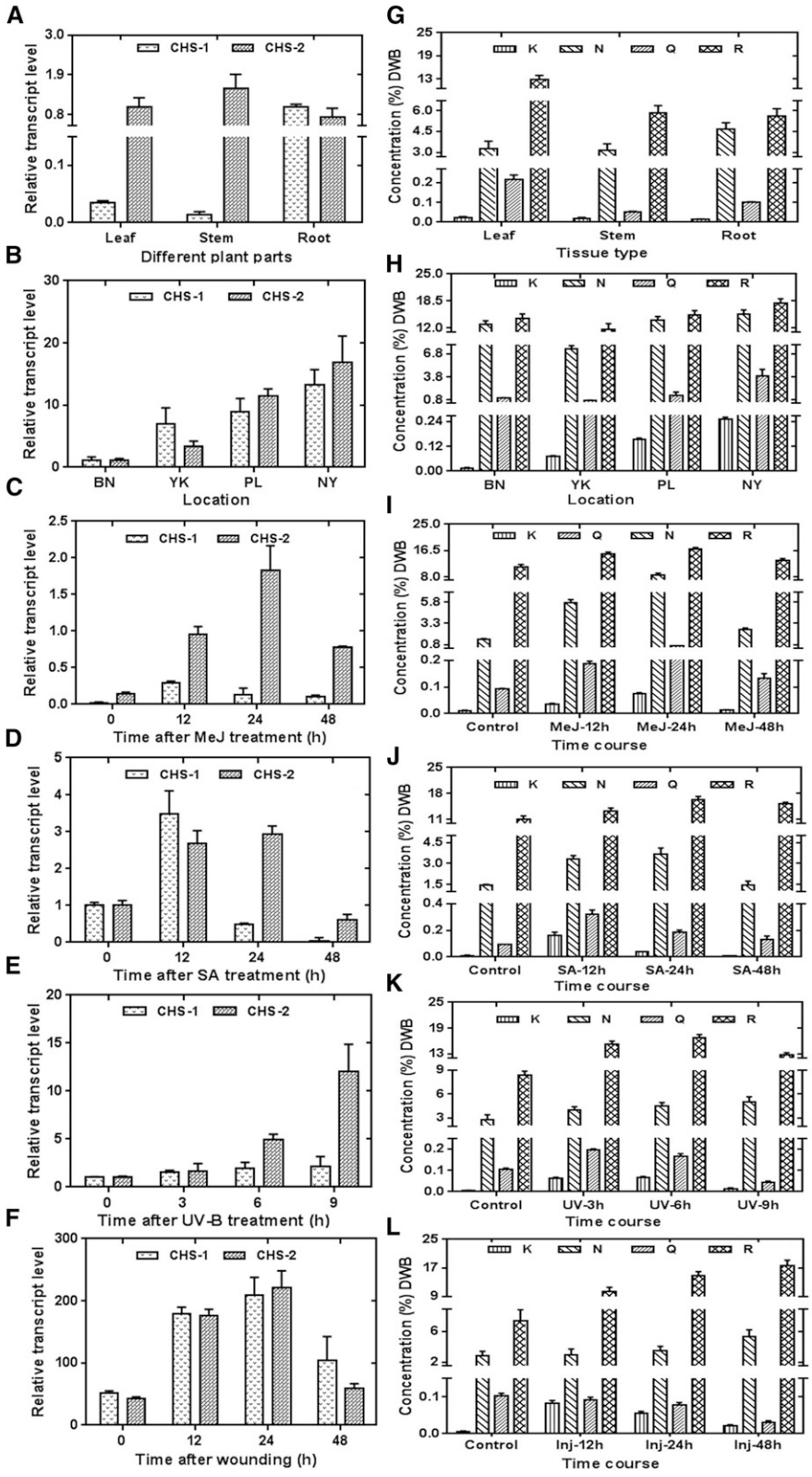

Figure 9. Real-time expression analysis of ReCHSs vis-à-vis flavonoid accumulation in *R. emodi*. A and G, Quantitative estimation of the relative expression of ReCHS1 and ReCHS2 (A) and the relative accumulation of flavonoids (G) in leaf, stem, and root tissues of *R. emodi*. B and H, Differential expression pattern of ReCHS1 and ReCHS2 (B) and total flavonoid content (H) in plant samples collected from four different geographic locations of northwestern Himalayas (1,600–4,500 m asl). C to F and I to L, Time-course expression characteristics of ReCHS1 and ReCHS2 and effects on flavonoid accumulation pattern in micropropagated *R. emodi* in response to elicitations by 0.1 mM MeJ (C and I), 0.1 mM SA (D and J), 1,500 μJ m^{-2} UV-B radiation (E and K), and wounding (F and L). The in vitro-raised cultures were precultured in Murashige and Skoog liquid medium for about 2 weeks, elicited, and further harvested after different time intervals, and the concentration of MeJ and SA for plant treatments was kept as 0.1 mM. β-Actin was used as an endogenous control to normalize the expression of ReCHSs. The data were compared and analyzed with one-way ANOVA using GraphPad Prism 6 software. Values are expressed as means ± SD, with SE values indicated by bars representing at least three replicates. Statistical significance was considered at $P < 0.001$. Locations are as follows: BN, Bonera Farm, Pulwama (33° 52′ 59′′ N, 74° 55′ 00′′ E; 1,630 m asl); YK, Yarikhah Farm, Srinagar (34° 04′ 797′′ N, 74° 26′ 448′′ E; 2,119 m asl); PL, Pense La Top, Ladakh (33° 51′ 08′′ N, 76° 21′ 57′′ E; 4,287 m asl); and NY, Nyoma Valley, Ladakh (33° 08′ 661′′ N, 78° 34′ 742′′ E; 4,415 m asl), K, Kaempferol; N, naringenin; Q, quercetin; and R, rutin. DWB, Dry weight basis.

shown that duplication and divergence of CHS sequences may occur before the speciation event in angiosperms (Beerhues and Liu, 2009). The equal length of the two ReCHSs excludes the possible occurrence of duplication events in them. However, our results indicate that synonymous/nonsynonymous mutation events

Table III. *Putative cis-acting regulatory elements identified in the promoters of ReCHS1 and ReCHS2 using PLACE (http://www.dna.affrc.go.jp/PLACE) and PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) databases*

cis-Element	Gene	Putative Function
CAAT box	*ReCHS-1, ReCHS-2*	Common cis-acting element in promoter and enhancer regions
GAG motif	*ReCHS-1*	Part of a light-responsive element
CGTCA motif	*ReCHS-1, ReCHS-2*	Cis-acting regulatory element involved in methyl jasmonate (MeJ) responsiveness
MBS	*ReCHS-2*	MYB-binding site involved in drought inducibility
G box	*ReCHS-1, ReCHS-2*	Cis-acting regulatory element involved in light responsiveness
MBSI	*ReCHS-2*	MYB-binding site involved in flavonoid biosynthetic gene regulation
MNF1	*ReCHS-1, ReCHS-2*	Light-responsive element
TGACG motif	*ReCHS-1*	Involved in transcriptional activation of several genes by auxin and/or SA
Skn-1 motif	*ReCHS-1, ReCHS-2*	Cis-acting regulatory element required for endosperm expression
O2 site	*ReCHS-2*	Cis-acting regulatory element involved in zein metabolism regulation
TC-rich repeats	*ReCHS-2*	Cis-acting element involved in defense and stress responsiveness
TATA box	*ReCHS-1, ReCHS-2*	Core promoter element around −30 of the transcription start
MRE	*ReCHS-1*	MYB-binding site involved in light responsiveness
TCA element	*ReCHS-2*	Cis-acting element involved in SA responsiveness
W box	*ReCHS-1, ReCHS-2*	Wound-responsive element
TCT motif	*ReCHS-2*	Part of a light-responsive element
TGA element	*ReCHS-2*	Auxin-responsive element

could have taken place over the period of evolution to generate the two paralogous members of ReCHS. In fact, it was shown recently that the rate of synonymous mutations was higher in the case of early-diverged CHSs than in the late ones (Han et al., 2014). Our observations of higher synonymous mutations in ReCHSs (Supplemental Table S1) are in agreement with this statement. Moreover, the isolated ReCHS paralogs were found to be intronless. Recent plant evolution has shown that intron losses astoundingly outnumbered the incidence of intron gain. For instance, Arabidopsis and rice (*Oryza sativa*) have reported 12.6 and 9.8 times more intron loss than gain, respectively (Roy and Penny, 2007). It also has been demonstrated that intronless genes exhibit high synonymous substitution rates (Yang et al., 2013).

The most popular and widely used method of proteomics, the bottom-up approach, was deployed to decipher the identity of ReCHS tryptic digests. In general, the sequence coverage in bottom-up proteomic analysis ranges from 5% to 70% (Chen and Pramanik, 2008). However, we were able to generate a better sequence coverage of ReCHS proteins. Moreover, the successful detection and identification of only two unique peptides is generally considered to be sufficient for protein identity (Ong and Mann, 2005). In this study, we have characterized five peptide sequences of ReCHS1 and six of ReCHS2 (Table I; Supplemental Figs. S3 and S4). This further supports the existence of two separate paralogous members of the small CHS gene family of *R. emodi*. With the advent of highly sensitive and advanced mass spectrometry analytical tools for proteomic characterization, it has become possible to complement the conventional western-blot techniques requiring antibody generation (Yates et al., 2009).

Even though the mass spectrometric identification of a specific protein may help in detecting its possible function, the eventual confirmation always rests on a clear biological experiment (Habermann et al., 2004). The catalytic activity of identified ReCHS enzymes demonstrated that they performed a typical CHS function by effectively catalyzing the synthesis of naringenin chalcone and/or naringenin. In our study, we were able to detect the presence of both isomeric forms of the flavonoid, naringenin chalcone and naringenin. Moreover, it was observed that the turnover rate of naringenin chalcone to chalcone was high. This is because we could spot the presence of small quantities of readily converting chalcone only after applying nonacidified reaction sample onto the LC-MS instrument immediately after the reaction was extracted (Fig. 5). Nonetheless, further experiments are needed to justify the statement. Mass spectrometry and fragmentation patterns of ReCHS products and their standards further validated their functionality, projecting them as suitable targets for future pathway-engineering endeavors. In addition, the kinetic characterization of ReCHSs demonstrated that the active site of ReCHS2 seems more flexible to nonphysiological substrates, as evident from its catalytic efficiency (Table II).

Furthermore, ReCHS1 was found to show a broad range of pH stability compared with that of ReCHS2, which was found to be active over a limited range of pH variance. It has been demonstrated that physicochemical constraints and evolutionary selective forces mold the kinetic parameters of enzymes regardless of the host organism. Evolutionary pressure also has been supposed to play an important role in shaping enzyme parameters, which in many cases could increase their catalytic efficiency toward natural substrates (Bar-Even et al., 2011). The higher catalytic efficiency (Table II) of ReCHS1 toward the main physiological substrate *p*-coumaroyl-CoA further validates our assumption.

Genetic redundancy is the characteristic feature of plant genomes, and nearly all plant genes examined so far are usually represented by small to large multigene families evolved over the course of evolution (Durbin et al., 2000). To explore the organization of CHS genes in the *R. emodi* genome, Southern-blot analysis was

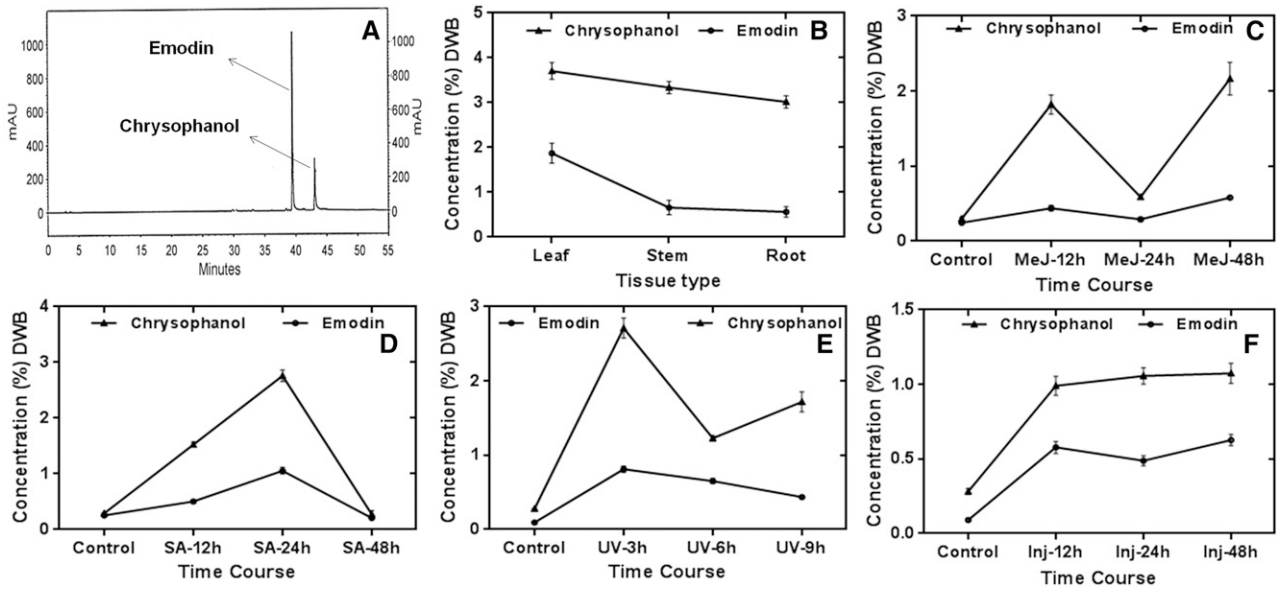

Figure 10. Quantification of major anthraquinones in *R. emodi* by HPLC analyses. A, Representative HPLC chromatogram of marker compounds (emodin and chrysophanol). B, Relative accumulation of major anthraquinones in leaf, stem, and root tissues of *R. emodi*. C to F, Time-course effects of elicitor treatments on the accumulation of anthraquinones in response to 0.1 mM MeJ (C), 0.1 mM SA (D), UV-B radiation (E), and wounding/injury (F) at different time intervals. In general, chrysophanol was found to accumulate in higher concentrations throughout the elicitations and at all harvesting periods. The data were compared and analyzed with one-way ANOVA using GraphPad Prism 6 software. Values are expressed as means ± SD, with SE values indicated by bars representing at least three replicates. The time-course accumulation of anthraquinones was statistically significant at $P < 0.001$. DWB, Dry weight basis; mAU, milli-absorbance unit.

performed. The results obtained suggest that the CHS family of *R. emodi* consists of at least two genes with a single copy for each of them. In other words, each of the CHS members exists in a single copy. Moreover, the substantial divergence in the 5′ flanking regions of the two paralogs excludes the possibility of their allelic nature. Furthermore, it is speculated that stress adaptations of plants result in the divergence of CHS genes into many isomers to combat the demand for flavonoid biosynthesis under stressful conditions. For example, stilbene synthase, which functions with the same substrates as CHS, is considered to have evolved independently several times in the course of evolution (Pang et al., 2005).

The results with tissue- and site-specific expression profiling demonstrate varied transcript levels of ReCHSs. In a broader context, the spatial expression pattern of ReCHSs was nearly in agreement with the content of flavonoids in respective tissues. Moreover, the slight variability in the accumulation of flavonoids with respect to the corroborative expression of ReCHSs seems to conform with the previous studies showing that flavonoids are capable of root-to-shoot and shoot-to-root movement, probably mediated by ABC-type transporters (Gutzeit and Ludwig-Müller, 2014). Besides spatial variation, the altitudinal effect also was evident in relation to flavonoid accumulation and the relative expression of ReCHSs. This trend of increasing flavonoid content in response to altitudinal gradients is well documented in earlier reports (Spitaler et al., 2006). The increased concentration of phenolic compounds and carotenoids with increasing altitude has been suggested as a response to high light intensity and escalating UV radiation. In particular, flavonoids have been reported as UV-B light-absorbing compounds (Jaakola et al., 2004).

The promoter-associated cis-acting regulatory elements along with their corresponding transcription factors constitute the transcriptional regulatory machinery subject to induction by diverse environmental and extracellular factors to facilitate the survival of plants in adverse conditions (Dhar et al., 2014). The promoter region of CHS has numerous cis-acting regulatory elements related to tissue specificity, stress, and phytohormones, among others, and they are induced by developmental and environmental factors such as UV light, wounding, and treatment with elicitors (Zhang et al., 2011). To gain insight into the regulatory mechanism of ReCHS promoter regions, the relative transcript levels were assayed in response to the application of endogenous and exogenous elicitors to understand their inducible/repressible nature and to substantiate these results with the recognized cis-acting elements (Table II; Supplemental Fig. S5, A and B). In general, the elicitations mediated by MeJ, SA, UV-B light, and wounding considerably up-regulated the relative transcript levels of ReCHSs and demonstrated the respective changes in the accumulation of flavonoid and

anthraquinone constituents. Additionally, Pearson's correlation coefficient analysis established a positive relation between the transcript levels of ReCHS1 and ReCHS2 and flavonoid and anthraquinone accumulation, respectively.

The signaling molecule MeJ has been recognized to modulate the biosynthesis of many secondary metabolites that play a vital role in the adaptation of plants to different kinds of stress conditions (Varadarajan et al., 2010). Treatment with MeJ augmented the expression of ReCHS genes, and the related transcript profiles were in complete agreement with similar changes in the content of flavonoids. This finding pointed out that ReCHS expression is subject to modulation by MeJ. Moreover, the exogenous treatment of plants with SA broadly displayed a transcript profile of paralogous ReCHS genes that was found to be in consonance with the increasing content of both flavonoids and anthraquinones. Additionally, several studies have shown that MeJ and SA are prospective candidates for the elicitation process leading to the transcriptional alteration of various genes that ultimately affect secondary metabolite biosynthesis (Ali et al., 2007).

Plants respond to high-intensity light and UV-B (290–320 nm) radiation either by exciting defense mechanisms or by stimulating repair mechanisms. The former usually involves the systemic accumulation of flavonoids, which play a vital role by acting as natural sunscreens (Brosché et al., 1999). The UV-B light exposure of plants resulted in well-coordinated increased expression of both paralogs, which paralleled the accumulation pattern of flavonoids only. In Arabidopsis, it has been shown that mutants deficient in phenylpropanoid compounds were considerably more susceptible to UV-B radiation than corresponding wild-type lines (Hectors et al., 2012). Furthermore, studies on parsley (*Petroselinum crispum*) cell cultures have shown that UV light is responsible for the induction of CHS expression by the de novo synthesis of the active enzyme (Dao et al., 2011). The herbivory or wounding induces metabolic changes that involve discerning the modulation of gene expression. The highest expression of ReCHS paralogs was observed upon subjecting plants to wounding with an increase in the metabolic content. The exposure of plants to wounding or UV-B irradiation results in the generation of reactive oxygen species or oxygen free radicals, which are quenched by flavonoids. Flavonoids are potent free radical-scavenging agents (Izaguirre et al., 2003). Moreover, UV-B radiation causing multiple cellular injuries has been shown to activate various signaling pathways, which ultimately result in the expression of genes responsible for the biosynthesis of important secondary metabolites (Jiao et al., 2015).

Promiscuity is often referred to as the process in which an enzyme coincidently catalyzes reactions other than the ones for which it has specifically evolved, or it may imply the increased enzymatic efficiency and product formation triggered by genic divergence. The homodimeric type III PKSs have been studied extensively for their substrate promiscuity and product diversity, making them an exceptional platform for the production of unnatural, novel polyketide scaffolds with promising biological activities (Bhan et al., 2015). The mutation-led promiscuous feature of enzymes is a rather common and widespread feature prevalent among many classes of enzymes, which shows a 10- to 10^6-fold increase in their activity. An enzyme can use the same active site for specific and promiscuous functions to carry out product biosynthesis by an altogether different mechanism utilizing unlike active site conformers (Adami, 2006). Enzymatic analysis has been shown to be an imperative tool to examine the functional divergence of a gene family (Des Marais and Rausher, 2008). ReCHS1 was seen to be involved primarily in flavonoid biosynthesis, owing to its higher specific activity (18.39 ± 0.47 nmol min^{-1} mg^{-1}) vis-à-vis substrate affinity with the main physiological substrate *p*-coumaroyl-CoA compared with that of ReCHS2 (10.37 ± 0.21 nmol min^{-1} mg^{-1}). The anthraquinones (like chrysophanol, a polyketidic anthraquinone) result from one or more intramolecular cyclization and aromatization events of the type III PKS-derived polyketide intermediate in the same active site. Although the metabolic pathway beyond the polyketide intermediate has yet to be elucidated, it has been postulated that the CHS family of type III PKSs may be involved in anthraquinone biosynthesis in plants (Austin and Noel, 2003). Furthermore, it has been contended that different active site geometries control different cyclization reactions in CHS, and the surface topology of the cyclization pocket directs polyketide folding and product formation. Alterations in the surface topology of the CHS cyclization pocket may affect the stereochemistry of the cyclization reaction and modulate product selectivity (Ferrer et al., 1999; Suh et al., 2000). In this study, it was observed that the conserved amino acid residues Ile-255 and Gly-257 were replaced by Leu and Ala in the cyclization pocket of ReCHS2. Increasing amino acid replacement rates are reported to be often examined in conjunction with shifts in enzyme function (Durbin et al., 2000). Thus, the differences in relative activity and enzyme efficiency as displayed by ReCHS paralogs could be a manifestation of the variation in their cyclization pocket. This finding may lend further credence to the promiscuity of ReCHS2.

Overall, we propose the possible promiscuous nature of ReCHSs characterized from *R. emodi* in relation to differential substrate selectivities and metabolite accumulation. These inferences are drawn from empirical experimental evidence based on enzymatic kinetic studies with five different substrates, real-time expression and metabolic profiling, and bottom-up proteomic and bioinformatic analyses of two evolutionarily divergent ReCHS paralogs.

MATERIALS AND METHODS

Chemicals

Malonyl-CoA, pure standards of major anthraquinones (emodin and chrysophanol) and flavonoids (kaempferol, quercetin, and rutin), DPBA for flavonoid staining, and the chemicals used for enzymatic digestion of proteins,

urea, dithiotreitol, iodoacetamide, and trypsin from bovine pancreas, were all purchased from Sigma-Aldrich. Naringenin chalcone was purchased from Apin Chemicals. Mass spectrometry-grade acetonitrile, water, and formic acid were purchased from Merck. All other reagents and chemicals used in this study were of HPLC or analytical grade.

Plant Selection and RNA Isolation

The plant material was originally collected from Pense La Top, Ladakh (33° 51′ 08′′ N, 76° 21′ 57′′ E; 4,287 m asl) as reported earlier by us (Pandith et al., 2014). The seeds collected were sown in earthen pots, and the germinated saplings maintained under greenhouse conditions at the Indian Institute of Integrative Medicine in Jammu, India (32° 44′ N longitude, 74° 55′ E latitude; 305 m asl), were used as source material for the establishment of the in vitro regeneration system and total RNA isolation. Total RNA was isolated (Ghawana et al., 2011) and incubated at 37°C for 30 min with DNase I (Fermentas) to eliminate the traces of genomic DNA. RNA quality was assessed by electrophoresis on 1% formaldehyde agarose gels and by determining the absorbance ratio ($A_{260/280}$) using a spectrophotometer (AstraAuriga).

cDNA Synthesis and Cloning of ReCHS1 and ReCHS2

For cDNA synthesis, 3 μg of DNase I-treated total RNA was reverse transcribed using the Revert-Aid premium reverse transcription kit (Fermentas) according to the manufacturer's instructions. Degenerate primers (Supplemental Table S3) were designed based on highly conserved regions of amino acid sequences of other plant CHSs retrieved from the GenBank database at the NCBI using BLASTN/BLASTX and ClustalW2 programs. Using cDNA as a template, reverse transcription-PCR for core amplification was carried out under the following cyclic conditions: one 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 50 s; and a final extension of 72°C for 10 min in a thermal cycler (Eppendorf). The selected amplicons were cloned separately into pTZ57R/T vector (Fermentas), transformed into an *Escherichia coli* host strain (DH5; Invitrogen), and further confirmed by sequencing analysis (ABI PRISM 3130XL genetic analyzer; Applied Biosystems). The sequenced core amplicons were used subsequently to design gene-specific primers to perform 5′ and 3′ RACE using the Gene Racer cDNA amplification kit according to the product manual (Invitrogen). By comparing and aligning the sequences of the core fragments, 5′ and 3′ RACE products, the full-length cDNAs of ReCHS1 and ReCHS2 were generated and subsequently amplified with full-length primers (FulCHS1_F/FulCHS1_R and FulCHS2_F/FulCHS2_R; Supplemental Table S3). The generated coding sequence fragments also were amplified from the intact DNA isolated from *Rheum emodi* to look for the possible existence of introns.

Bioinformatic Analysis

The complete coding sequence fragments obtained were translated using the Translate tool, and secondary structures predicted by SOPMA and the properties of deduced amino acid sequences were estimated using ProtParam and Compute pI/Mw programs. The presence of signal peptides was evaluated using SignalP 4.1, and TMHMM version 2.0 was used to predict the number of transmembrane helices. The identification of functionally and structurally important residues in protein sequences was done using ConSeq and ConSurf programs (Ashkenazy et al., 2010). The 3D structures of proteins were determined by I-TASSER and Phyre2 software using the crystal structure of alfalfa (*Medicago sativa*) CHS (PDB code c1cmla) as a template. The standard molecular viewer, PyMOL Molecular Graphics System, version 1.7.4 (Schrödinger; http://www.pymol.org/), was used extensively in the preparation of 3D protein structures. The ligand-binding sites were predicted by the 3DLigandSite server. Subcellular localization analysis was performed using TargetP 1.1 (Emanuelsson et al., 2007). The 3D structures of ReCHS1 and ReCHS2 were compared with that of the model structure of alfalfa CHS (PDB code c1cmla) using the Pairwise Alignment Tool of the FATCAT program (Ye and Godzik, 2003). Moreover, the amino acid sequences were aligned using the Clustal Omega Multiple Sequence Alignment tool, and the phylogenetic analysis was performed using the ClustalW program and MEGA6 software based on the neighbor-joining method with 1,000 bootstrap replicates to obtain confidence levels with the branches. In addition, the Synonymous Nonsynonymous Analysis Program (SNAP version 2.1.1), based on a modified method of Nei and Gojobori (1986), was used to calculate synonymous and nonsynonymous substitution rates in ReCHS sequences.

Plasmid Construction, Heterologous Expression, and Purification of Recombinant ReCHSs

The ORFs of ReCHS1 and ReCHS2 were tailored using sense and antisense primers by the addition of *Bam*HI and *Eco*RI restriction sites upstream and downstream to start and stop codons, respectively. The resulting CHSs were cloned and excised from pJET vector (Fermentas) with *Bam*HI and *Eco*RI and reconfirmed by sequencing before subcloning into the respective restriction sites of predigested and purified bacterial expression vector pGEX-4T-2. The cloned CHS superfamily proteins were expressed as fusion proteins with a GST tag at the N terminus of the expression vector. Heterologous expression of the recombinant proteins was carried out as described earlier (Bhat et al., 2014). For purification, *E. coli* BL21 (DE3) cells transformed with respective expression plasmids of ReCHS1 and ReCHS2 were grown in Luria-Bertani medium with the desired antibiotic at 37°C until A_{600} = 0.4 to 0.6, and expression was induced with 0.6 mM IPTG. After an induction period of 6 to 8 h at 30°C, cells were harvested by centrifugation (6,000*g*, 4°C, and 10 min; Eppendorf), resuspended in 1× PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 10 mM KH_2PO_4, pH 7.3), lysed using 20 mM dithiotreitol and 0.2 mg mL^{-1} lysozyme, and further incubated on ice for 30 min to effect lysis. The culture was briefly sonicated (3 × 30 s) using a probe sonicator (Sartorius) and incubated on ice for 30 min with 1% (v/v) Triton X-100. The soluble fraction was recovered by centrifugation (12,000*g*, 4°C, and 15 min) and further incubated with processed glutathione-Sepharose 4B resin (1 mL L^{-1} culture; GE Healthcare) at 4°C overnight. The Sepharose beads were washed four to five times with 10 bed volumes of 1× PBS and incubated with thrombin protease (less than 10 cleavage units mL^{-1}) at 24°C for 12 to 16 h with gentle shaking to remove the GST moiety. The suspension was centrifuged for 3 min at 3,000 rpm, and the supernatant was incubated with benzamidine-Sepharose (10 μL unit^{-1} thrombin protease) for 30 min at 22°C to 24°C to remove the thrombin. The cleaved protein samples were denatured and analyzed by 10% SDS-PAGE, and the concentration was measured using the Bradford colorimetric protein assay.

Bottom-Up Proteomic Characterization of ReCHSs

Tryptic Digestion

The purified ReCHS proteins were characterized using the bottom-up proteomic approach of mass spectrometry involving the enzymatic digestion of proteins before subjecting them to mass spectrometry analysis. The isolated ReCHS proteins were subjected to complete in-solution tryptic digestion including denaturation, reduction, and alkylation steps, as per the manufacturer's instructions (Sigma-Aldrich).

LC-MS, ESI-Mass Spectrometry, and Ultra-High-Definition/Quadrupole Time-of-Flight Mass Spectrometry

The Agilent 1290 HPLC system, coupled with the Agilent 6540 ultra-high-definition/quadrupole time-of-flight analyzer and equipped with RP-18e (Merck Purospher; 4.6 × 250 mm), was used for the analysis of tryptic digests of ReCHS1 and ReCHS2 protein samples. A binary gradient of water with 0.1% (v/v) formic acid (solvent A) and acetonitrile (solvent B) was applied as 25 min 50/50 (%A/%B), 35 min 10/90, 37 min 95/5, and 40 min 95/5, with a flow rate of 400 μL min^{-1}. The injection volume of the sample was 20 μL. The instrumental parameters for proteomic analysis were set as follows: ionization mode, dual electrospray ionization mode; N_2 gas flow, 12 L min^{-1}; gas temperature, 300°C; nebulizer pressure, 45 p.s.i.; and capillary voltage, 3.5 kV. The mass spectrometer was set at an m/z range of 100 to 3,000 atomic mass units in positive ion mode. The collision energy (CE) was set according to the charge state of the ion on the basis of the calculation formula CE = (slope × 0.01 [m/z] + offset). The slope and offset were 4 and 3 for singly charged ions, 5 and 4 for doubly charged ions, 5.5 and 5 for triply charged ions, and 6.5 and 6 for higher charged ions, respectively. The data were analyzed with Agilent Mass Hunter BioConfirm version B.06.00 software.

In Vitro Enzyme Assay and Steady-State Kinetics

The activities of purified ReCHSs were determined separately by monitoring product formation using LC-MS analysis. The respective standard reaction mixtures contained purified enzyme (20 μg), 60 μM starter-CoA (*p*-coumaroyl-CoA), and 150 μM malonyl-CoA in a 100-μL reaction of 0.1 M potassium

phosphate (pH 7), 1 mM EDTA, and 10% glycerol. The reactions were incubated at 30°C for 1 h, quenched by acidification (10 μL of 20% HCl), and further extracted with ethyl acetate (3 × 200 μL) to collect the soluble fraction. The extracts were evaporated to dryness and redissolved in methanol. Naringenin and naringenin chalcone were used as reference compounds for the identification of reaction products of the purified ReCHS1 and ReCHS2 proteins. Some of the reactions from both samples were extracted directly without quenching, and products from both stopped and unstopped reactions of each of the purified proteins were initially applied to thin-layer chromatography plates (30% ethyl acetate in hexane) to check the activity of enzymes. The extracts were subjected to LC-MS analysis to confirm and quantify naringenin and naringenin chalcone produced.

The steady-state kinetic parameters for five different substrates were determined from initial velocity measurements where product formation was linear over the monitored time periods, using standard assay conditions with a fixed malonyl-CoA concentration (120 μM) and varied starter-CoA concentrations (10–250 μM). The kinetic constants K_m and V_{max} were calculated with nonlinear regression analysis using GraphPad Prism 6 software. The relative activities of purified ReCHSs also were observed using different starter esters.

Effect of pH on Enzyme Activity

In general, the CHSs are known for their pH-dependent activity. It is known that naringenin chalcone slowly gets converted to its cyclized isomer naringenin in the aqueous state and upon acidification under in vitro conditions. Therefore, the impact of pH on ReCHS activity was analyzed by determining the formation of naringenin as the major product. Different buffer systems were used to check the activity of purified proteins over a wide range of pH values (pH 5–9). The buffer systems include citrate buffer (pH 3–6.2), potassium phosphate buffer (pH 5.8–8), and 0.1 M Tris-HCl for pH 8.5 and 9. The reaction setup was the same as above, with the exception that all reactions were quenched to study the activities of ReCHS1 and ReCHS2 referenced only to the formation of naringenin.

HPLC-ESI-MS/MS Analysis

Stock solutions (1 mg mL^{-1}) of naringenin and naringenin chalcone were freshly prepared in methanol, filter sterilized with 0.25-μm membrane filters (Millipore), and stored at 4°C until further use. Standard working solutions were obtained by making appropriate dilutions of stock solutions for the preparation of a six-point calibration curve. The analyses were performed using an Agilent 1260 Infinity HPLC system equipped with 1260VL infinity quaternary pumps, an autosampler, and a thermostat compartment. The samples were separated on a Purospher STAR RP-18e column (100 × 4.6 mm, 5-μm particle size). Mobile phases consisted of 0.1% (v/v) formic acid in water (eluent A) and acetonitrile with 0.1% (v/v) formic acid (eluent B). A gradient program was used as follows: 0 to 10 min, 50% to 80% B; 10 to 15 min, 80% B; 15 to 17 min, 80% to 50% B; and 17 to 20 min, 50% B. The flow rate was adjusted to 0.3 mL min^{-1}, and column temperature was maintained at 30°C. Triple quadrupole MS/MS was carried out on an Agilent 6410 tandem triple quadrupole mass spectrometer equipped with an ESI ion source operating in both positive and negative ion modes. The ESI source was operated in positive ionization mode, and quantification was performed in MRM mode. The mass spectrometry parameters optimized were capillary voltage of 4.0 kV and gas temperature of 300°C. Nitrogen was used as the desolvation gas at the rate of 12 L min^{-1}, and nebulizer pressure was maintained at 50 p.s.i. Nitrogen also was used as the collision gas. All data were collected in the centroid mode and acquired and further processed using Mass Hunter work station software (Agilent). Several liquid chromatography parameters were optimized to obtain better separation and higher sensitivity with reduced analysis time. The high-quality separation was achieved with the Purospher STAR RP-18e column (100 × 4.6 mm, 5-μm particle size). Mass spectrometry scan mode conditions were optimized using the reference compounds, and higher sensitivity and clear mass spectra were observed in analyses conducted in the positive ion mode. In positive ion mode, quasimolecular ions ($[M+H]^+$) of naringenin and naringenin chalcone were generated, whose product ions were high with good specificity. The optimized fragmentor voltage and collision energy for both naringenin and naringenin chalcone were 130 V and 17 eV, respectively. Quantification was performed in MRM mode, having the ion transitions for naringenin and naringenin chalcone as m/z 273/153 and 273/147, respectively. The developed method showed a 14.9-min retention time for naringenin and 13.8 min for naringenin chalcone (Fig. 2). Compounds were identified by comparison of molecular ions, fragmented ions (MRM), and retention times with those of the standard compounds.

Southern-Blot Hybridization of Genomic DNA

Genomic DNA was isolated from the leaves of *R. emodi* using the DNeasy plant mini kit (Qiagen). Aliquots of the isolated and purified DNA (greater than 20 μg) were digested overnight at 37°C with restriction endonucleases *Bam*HI (noncutter) and *Nco*I (single cutter) for ReCHS1 and *Eco*RI (noncutter) and *Xma*I (single cutter) for ReCHS2. The digested DNA samples were separated by electrophoresis on 0.8% (w/v) agarose gels in Tris-borate/EDTA buffer and then transferred onto a positively charged nylon membrane (Roche) as described earlier (Rana et al., 2013). The probes corresponding to ReCHS1 and ReCHS2 were synthesized separately for hybridization by reverse transcription-PCR using the purified coding sequences of respective full-length cDNAs as templates. Probe labeling, hybridization, blocking, washing, and signal detection steps were performed according to the manufacturer's protocol as defined in the DIG-DNA labeling and detection kit (Roche).

In Vitro Regeneration, and Tissue-Specific and Site-Specific Gene Expression Analyses

The germinated saplings from greenhouse pots were used for the establishment of axenic cultures of *R. emodi*. Cultures established in Murashige and Skoog medium supplemented with 3% (w/v) Suc, 3.7 μM IBA, and 8.9 μM BAP were kept under a 16-h photoperiod with a light intensity of 25 to 30 μE m^{-2} s^{-1} provided by 40-W cool-white fluorescent lamps (Philips). Cultures were maintained at 25°C ± 1°C and a relative humidity of 50% to 60% (data not shown). The tissue-specific as well as spatial expression profiling was done by quantitative real-time PCR analysis. Total RNA was isolated from different parts (leaf, stem, and root) of plants and from the plant samples (leaf tissues) collected from four different geographic locations of Kashmir Himalayas. For each sample, DNase-treated RNA (3 μg) was reverse transcribed using the iScript cDNA synthesis kit (Bio-Rad) according to the manufacturer's instructions. SYBR-based chemistry using SYBR Premix Ex Taq (Takara) was applied in the ABI StepOne real-time quantitative PCR system (Applied Biosystems) to run the PCR. The respective PCRs of 10 μL included 0.5 μL of cDNA as template, 0.2 μM each primer (Supplemental Table S3), 5 μL of SYBR Premix Ex Taq, and Milli-Q water (Millipore) to make up the final volume. The reaction thermo profile was followed as recommended by the manufacturer: holding stage of one cycle at 95°C for 10 min; 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 primer design was done by Primer Express version 3.0 (Applied Biosystems) and was further validated by a dissociation curve (observation of a single peak for each primer pair). All samples were run in triplicate, and a housekeeping gene, *β-ACTIN*, amplified with Actin_F and Actin_R primers (Supplemental Table S3), was used as an endogenous control to normalize the expression of ReCHSs. The amplification of the target genes was monitored every cycle by SYBR Green fluorescence. The real-time amplification data were exported to Microsoft Excel, further analyzed by the method of Livak and Schmittgen (2001), and expressed as normalized relative expression levels ($2^{-\Delta\Delta CT}$) of the respective genes in various samples.

Amplification of Promoter Regions of ReCHSs

ReCHS promoter sequences were determined by the genome-walking method using the GenomeWalker Universal Kit (Clontech). Primers were designed and experiments were performed according to the kit protocol (Supplemental Table S3). GenomeWalker DNA libraries were constructed according to the user manual (Clontech). Genomic DNA was isolated from in vitro-raised plants of *R. emodi* using the DNeasy plant mini kit (Qiagen) according to the manufacturer's instruction manual. Isolated DNA was digested in four separate aliquots, 5 μg each, for the construction of GenomeWalker DNA libraries with the use of four different blunt-end-generating restriction endonucleases (*Dra*I, *Pvu*II, *Eco*RV, and *Stu*I). Each set of digested genomic DNA samples was purified and ligated to the GenomeWalker AP adaptor (provided with the kit) independently to generate four adapter-ligated libraries to be used as templates in PCRs for DNA

walking. To clone the promoter regions of ReCHS1 and ReCHS2, two-round PCR was performed using the adapter (provided with the kit) and gene-specific primers (Supplemental Table S3) as per the manufacturer's instruction manual. The putative cis-acting elements that extended upstream to the start codons of ReCHS1 and ReCHS2 were identified by searching the PlantCare and PLACE databases.

Plant Treatments for Elicitor Assays

To examine the variation in the accumulation of ReCHS transcript levels upon elicitor treatments, in vitro-raised cultures of *R. emodi* were precultured in Murashige and Skoog liquid medium for about 2 weeks. The same medium was used for exogenous supplements of elicitors, as liquid cultures tend to provide a better response than static ones. The stabilized plant cultures were subjected to different elicitor treatments in congruence with the promoter motifs identified. Elicitor treatments of MeJ (0.1 mM) and SA (0.1 mM) for 12, 24, and 48 h were given, and tissue harvesting was done at defined time periods for RNA, flavonoid, and anthraquinone extraction. Some of the cultures were subjected to wounding, wherein plants were injured/incised on nodal sections and leaves to study the effect on ReCHS expression level. The tissue harvesting was done after 12, 24, and 48 h. Furthermore, the adapted cultures also were exposed to UV-B light stress, wherein the cultures were kept in a dark closed chamber and exposed for 3, 6, and 9 h to 1,500 $\mu J\ m^{-2}$ UV-B irradiation. The untreated plant cultures were kept as controls, and the control culture for UV-B light was kept in a dark closed chamber. The RNA samples from all the treated and control cultures were reverse transcribed as discussed above. The effects of elicitor treatments on the expression analysis of ReCHS1 and ReCHS2 were studied using real-time PCR analysis with the same parameters as mentioned above.

Extraction and Quantification of Total Flavonoids and Anthraquinones by HPLC

The plant samples were dried under a gentle air stream (temperature of 25°C ± 2°C and relative humidity of 65% ± 5%) and pulverized to fine powder using a mortar and pestle. The powdered samples were serially extracted (3 × 100 mL) with DCM:methanol in the ratio of 1:1 (v/v). The extractions were done at room temperature over a period of 72 h (24 × 3), and every time, fresh solvents were used for the left out marc. The filtrates were combined and filtered through Whatman No. 1 paper, and solvents were removed at 45°C under reduced pressure using a rotary evaporator (Sigma-Aldrich) to yield the extract. The stock solutions (1 $mg\ mL^{-1}$) of flavonoids along with extracts were freshly dissolved in methanol and filter sterilized with 0.25-μm membrane filters (Millipore). An HPLC device (Shimadzu CLASS-VP V 6.14 SPI model) equipped with an RP-18e column (Merck; 5 μm, 4.6 × 250 nm), a photo-diode array detector (SPD-M10A VP model), and a pump (LC-10AT VP model) was used for the analysis of both flavonoids and anthraquinones. The analysis and determination of major anthraquinone constituents in plant samples were done as described previously (Pandith et al., 2014). However, a standard method (Samappito et al., 2002) with slight modifications with the same operating system was used for the determination of flavonoid constituents (kaempferol, naringenin, quercetin, and rutin) in the plant samples. The solvent system was 97.8% (v/v) water, 2% CH_3CN, and 0.2% H_3PO_4 (A) and 97.8% (v/v) CH_3CN, 2% water, and 0.2% H_3PO_4 (B), with a gradient elution of 0 to 30 min, 20% B; 30 to 35 min, 45% B; 35 to 38 min, 55% B; 38 to 40 min, 55% B; and 40 to 45 min, 20% B, at a flow rate of 0.6 mL min^{-1}. The injection volume of the sample was 20 μL, and the column temperature was 30°C. The identification, detection, and quantification of flavonoids were done on the basis of the retention times of reference compounds under a specific set of column operating conditions. Elution positions were established with authentic samples and by comparison with literature data. Relative contents of different flavonoid constituents were determined and expressed as percentage peak area. The flavonoid content was monitored at 370 nm.

In Situ Flavonoid Staining with DPBA

The plant samples subjected to flavonoid staining were the same as used for real-time expression analysis and for the determination of total flavonoid (and anthraquinone) content. The flavonoids were visualized using DPBA, a flavonoid-specific fluorescent dye (also known as Natural Product Reagent A; Roth 9920) by adopting a method used by Stracke et al. (2010). The different plant tissues and elicitor-treated harvested samples were washed with 1× PBS and then submerged in a freshly prepared aqueous solution of 0.25% (w/v) DPBA and 0.02% (v/v) Triton X-100 for at least 2 h. The treated samples were again washed two to three times with 1× PBS buffer and mounted onto slides for microscopy. Fluorescence was visualized with a stereomicroscope (Nikon SMZ1000) equipped with a GFP filter unit with an excitation wavelength of 480/40 nm and a 535/50-nm barrier filter. The leaf and intermodal segments (2–3 cm above the basal portion of the plant) were used for imaging of elicitor-treated samples. The photographs were taken both in bright fields and fluorescence fields with the same parameters set for imaging of all samples. Untreated samples of different plant tissues used as negative controls also were subjected to photography in both fluorescence and bright imaging fields. Moreover, samples with and without DPBA staining were used to configure microscopic parameters to make sure that fluorescence was from DPBA-stained flavonols.

Statistical Analysis

All experiments were analyzed with at least three replicates. The values of flavonoid and anthraquinone contents and also gene expression analyses are expressed as means ± SD. Statistical analyses were carried out using GraphPad Prism 6 software (GraphPad Software) by one-way ANOVA, and statistical significance was considered at $P < 0.001$. Furthermore, Pearson's correlation test was used to correlate ReCHS transcript levels with the metabolite accumulation, and the analyses were conducted using GraphPad Prism 6 software. Mean values of major flavonoid (kaempferol, quercetin, naringenin, and rutin) and anthraquinone (emodin and chrysophanol) constituents were cumulatively considered for correlation analyses.

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under the following accession numbers: ReCHS1 (KF850684) and ReCHS2 (KC822472).

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Conserved residue prediction for ReCHSs.

Supplemental Figure S2. Predicted 3D models and ligand-binding sites of ReCHSs.

Supplemental Figure S3. Liquid chromatography-MS/MS of tryptic digest of ReCHS1.

Supplemental Figure S4. Liquid chromatography-MS/MS of tryptic digest of ReCHS2.

Supplemental Figure S5. Nucleotide sequences of ReCHS1 and ReCHS2 gene promoters.

Supplemental Figure S6. Flavonoid accumulation in different tissues of *R. emodi* as determined by DPBA staining.

Supplemental Table S1. Synonymous and nonsynonymous mutations of ReCHSs.

Supplemental Table S2. Major flavonoid and anthraquinone contents from different plant tissues/samples.

Supplemental Table S3. List of primers used in the study.

Supplemental Table S4. Correlation of mRNA levels with flavonoid and anthraquinone contents in different tissues of *R. emodi*.

Supplemental Table S5. Correlation of mRNA levels with flavonoid and anthraquinone contents in plant samples of *R. emodi* from different geographic locations (1,600–4,500 m asl).

Supplemental Table S6. Correlation of mRNA levels with flavonoid and anthraquinone contents in MeJ-treated plants of *R. emodi*.

Supplemental Table S7. Correlation of mRNA levels with flavonoid and anthraquinone contents in SA-treated plants of *R. emodi*.

Supplemental Table S8. Correlation of mRNA levels with flavonoid and anthraquinone contents in UV light-treated plants of *R. emodi*.

Supplemental Table S9. Correlation of mRNA levels with flavonoid and anthraquinone contents in wounded plants of *R. emodi*.

ACKNOWLEDGMENTS

We thank Werner Heller (Institute of Biochemical Plant Pathology, Neuherberg, Germany) for providing *p*-coumaroyl-CoA; Dr. S.J.S. Flora (Defence Research and Development Establishment, Gwalior, India) for the gift of naringenin; two anonymous reviewers for their critical reviews and helpful comments that improved the article; and Rajneesh Anand at CSIR-IIIM for facilitating HPLC and HRMS analyses. This article represents institutional communication IIIM/1866/2015.

Received March 25, 2016; accepted June 6, 2016; published June 7, 2016.

LITERATURE CITED

Abe I, Morita H (2010) Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. Nat Prod Rep **27:** 809–838

Adami C (2006) Evolution: reducible complexity. Science **312:** 61–63

Ali MB, Hahn EJ, Paek KY (2007) Methyl jasmonate and salicylic acid induced oxidative stress and accumulation of phenolics in Panax ginseng bioreactor root suspension cultures. Molecules **12:** 607–621

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:** W529–W533

Austin MB, Noel JP (2003) The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep **20:** 79–110

Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R (2011) The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry **50:** 4402–4410

Beerhues L, Liu B (2009) Biosynthesis of biphenyls and benzophenones: evolution of benzoic acid-specific type III polyketide synthases in plants. Phytochemistry **70:** 1719–1727

Bhan N, Cress BF, Linhardt RJ, Koffas M (2015) Expanding the chemical space of polyketides through structure-guided mutagenesis of Vitis vinifera stilbene synthase. Biochimie **115:** 136–143

Bhat WW, Rana S, Dhar N, Razdan S, Pandith SA, Vishwakarma R, Lattoo SK (2014) An inducible NADPH-cytochrome P450 reductase from Picrorhiza kurrooa: an imperative redox partner of cytochrome P450 enzymes. Funct Integr Genomics **14:** 381–399

Brosché M, Fant C, Bergkvist SW, Strid H, Svensk A, Olsson O, Strid Å (1999) Molecular markers for UV-B stress in plants: alteration of the expression of four classes of genes in Pisum sativum and the formation of high molecular mass RNA adducts. Biochim Biophys Acta **1447:** 185–198

Chen G, Pramanik BN (2008) LC-MS for protein characterization: current capabilities and future trends. Expert Rev Proteomics **5:** 435–444

Dao TT, Linthorst HJ, Verpoorte R (2011) Chalcone synthase and its functions in plant resistance. Phytochem Rev **10:** 397–412

Deng X, Bashandy H, Ainasoja M, Kontturi J, Pietiäinen M, Laitinen RA, Albert VA, Valkonen JP, Elomaa P, Teeri TH (2014) Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of Gerbera hybrida. New Phytol **201:** 1469–1483

Des Marais DL, Rausher MD (2008) Escape from adaptive conflict after duplication in an anthocyanin pathway gene. Nature **454:** 762–765

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:** 17249–17267

Dixon RA, Lamb CJ, Masoud S, Sewalt VJ, Paiva NL (1996) Metabolic engineering: prospects for crop improvement through the genetic manipulation of phenylpropanoid biosynthesis and defense responses. A review. Gene **179:** 61–71

Domon B, Costello CE (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj J **5:** 397–409

Durbin ML, McCaig B, Clegg MT (2000) Molecular evolution of the chalcone synthase multigene family in the morning glory genome. Plant Mol Biol **42:** 79–92

Ehrenberg R (2015) Engineered yeast paves way for home-brew heroin. Nature **521:** 267–268

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

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 Crops Prod **50:** 112–117

Ferrer JL, 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:** 775–784

Ghawana S, Paul A, Kumar H, Kumar A, Singh H, Bhardwaj PK, Rani A, Singh RS, Raizada J, Singh K, et al (2011) An RNA isolation system for plant tissues rich in secondary metabolites. BMC Res Notes **4:** 85

Guo XD, Ma YJ, Parry J, Gao JM, Yu LL, Wang M (2011) Phenolics content and antioxidant activity of tartary buckwheat from different locations. Molecules **16:** 9850–9867

Gutzeit HO, Ludwig-Müller J (2014) Plant Natural Products: Synthesis, Biological Functions and Practical Applications. John Wiley & Sons

Habermann B, Oegema J, Sunyaev S, Shevchenko A (2004) The power and the limitations of cross-species protein identification by mass spectrometry-driven sequence similarity searches. Mol Cell Proteomics **3:** 238–249

Han Y, Zhao W, Wang Z, Zhu J, Liu Q (2014) Molecular evolution and sequence divergence of plant chalcone synthase and chalcone synthase-like genes. Genetica **142:** 215–225

Han YY, Ming F, Wang W, Wang JW, Ye MM, Shen DL (2006) Molecular evolution and functional specialization of chalcone synthase superfamily from Phalaenopsis orchid. Genetica **128:** 429–438

Hectors K, van Oevelen S, Guisez Y, Prinsen E, Jansen MA (2012) The phytohormone auxin is a component of the regulatory system that controls UV-mediated accumulation of flavonoids and UV-induced morphogenesis. Physiol Plant **145:** 594–603

Izaguirre MM, Scopel AL, Baldwin IT, Ballaré CL (2003) Convergent responses to stress: solar ultraviolet-B radiation and Manduca sexta herbivory elicit overlapping transcriptional responses in field-grown plants of *Nicotiana longiflora*. Plant Physiol **132:** 1755–1767

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

Jenkins GI, Long JC, Wade HK, Shenton MR, Bibikova TN (2001) UV and blue light signalling: pathways regulating chalcone synthase gene expression in Arabidopsis. New Phytol **151:** 121–131

Jez JM, Austin MB, Ferrer J, Bowman ME, Schröder J, Noel JP (2000) Structural control of polyketide formation in plant-specific polyketide synthases. Chem Biol **7:** 919–930

Jiang C, Schommer CK, Kim SY, Suh DY (2006) Cloning and characterization of chalcone synthase from the moss, Physcomitrella patens. Phytochemistry **67:** 2531–2540

Jiang W, Yin Q, Wu R, Zheng G, Liu J, Dixon RA, Pang Y (2015) Role of a chalcone isomerase-like protein in flavonoid biosynthesis in Arabidopsis thaliana. J Exp Bot **66:** 7165–7179

Jiao J, Gai QY, Wang W, Luo M, Gu CB, Fu YJ, Ma W (2015) Ultraviolet radiation-elicited enhancement of isoflavonoid accumulation, biosynthetic gene expression, and antioxidant activity in Astragalus membranaceus hairy root cultures. J Agric Food Chem **63:** 8216–8224

Lau W, Sattely ES (2015) Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. Science **349:** 1224–1228

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

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. Methods **25:** 402–408

Luo Y, Li BZ, Liu D, Zhang L, Chen Y, Jia B, Zeng BX, Zhao H, Yuan YJ (2015) Engineered biosynthesis of natural products in heterologous hosts. Chem Soc Rev **44:** 5265–5290

Mounet F, Moing A, Garcia V, Petit J, Maucourt M, Deborde C, Bernillon S, Le Gall G, Colquhoun I, Defernez M, et al (2009) Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. Plant Physiol **149:** 1505–1528

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol **3:** 418–426

Ong SE, Mann M (2005) Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol **1:** 252–262

Pandey A, Misra P, Choudhary D, Yadav R, Goel R, Bhambhani S, Sanyal I, Trivedi R, Trivedi PK (2015) AtMYB12 expression in tomato leads to large scale differential modulation in transcriptome and flavonoid content in leaf and fruit tissues. Sci Rep **5:** 12412

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

Pang Y, Shen G, Wu W, Liu X, Lin J, Tan F, Sun X, Tang K (2005) Characterization and expression of chalcone synthase gene from Ginkgo biloba. Plant Sci **168:** 1525–1531

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:** 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:** 982–987

Rokaya MB, Münzbergová Z, Timsina B, Bhattarai KR (2012) Rheum australe D. Don: a review of its botany, ethnobotany, phytochemistry and pharmacology. J Ethnopharmacol **141:** 761–774

Roy SW, Penny D (2007) Patterns of intron loss and gain in plants: intron loss-dominated evolution and genome-wide comparison of O. sativa and A. thaliana. Mol Biol Evol **24:** 171–181

Samappito S, Page J, Schmidt J, De-Eknamkul W, Kutchan TM (2002) Molecular characterization of root-specific chalcone synthases from Cassia alata. Planta **216:** 64–71

Spitaler R, Schlorhaufer PD, Ellmerer EP, Merfort I, Bortenschlager S, Stuppner H, Zidorn C (2006) Altitudinal variation of secondary metabolite profiles in flowering heads of Arnica montana cv. ARBO. Phytochemistry **67:** 409–417

Stafford HA (1990) Flavonoid Metabolism. CRC Press, Boca Raton, FL

Stracke R, Jahns O, Keck M, Tohge T, Niehaus K, Fernie AR, Weisshaar B (2010) Analysis of PRODUCTION OF FLAVONOL GLYCOSIDES-dependent flavonol glycoside accumulation in Arabidopsis thaliana plants reveals MYB11-, MYB12- and MYB111-independent flavonol glycoside accumulation. New Phytol **188:** 985–1000

Suh DY, Fukuma K, Kagami J, Yamazaki Y, Shibuya M, Ebizuka Y, Sankawa U (2000) Identification of amino acid residues important in the cyclization reactions of chalcone and stilbene synthases. Biochem J **350:** 229–235

Tosetti R, Tardelli F, Tadiello A, Zaffalon V, Giorgi FM, Guidi L, Trainotti L, Bonghi C, Tonutti P (2014) Molecular and biochemical responses to wounding in mesocarp of ripe peach (Prunus persica L. Batsch) fruit. Postharvest Biol Technol **90:** 40–51

Tossi V, Lombardo C, Cassia R, Lamattina L (2012) Nitric oxide and flavonoids are systemically induced by UV-B in maize leaves. Plant Sci **193-194:** 103–109

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

Varadarajan J, Guilleminot J, Saint-Jore-Dupas C, Piégu B, Chabouté ME, Gomord V, Coolbaugh RC, Devic M, Delorme V (2010) ATR3 encodes a diflavin reductase essential for Arabidopsis embryo development. New Phytol **187:** 67–82

Walter JM (2014) Engineering yeast to produce artemisinic acid for antimalarial drugs. *In* Annual Meeting and Exhibition 2014 (July 20–24, 2014). https://sim.confex.com/sim/2014/webprogram/Paper27488.html

Wang SY, Bowman L, Ding M (2008) Methyl jasmonate enhances antioxidant activity and flavonoid content in blackberries (Rubus sp.) and promotes antiproliferation of human cancer cells. Food Chem **107:** 1261–1269

Wu S, Chappell J (2008) Metabolic engineering of natural products in plants: tools of the trade and challenges for the future. Curr Opin Biotechnol **19:** 145–152

Yang YF, Zhu T, Niu DK (2013) Association of intron loss with high mutation rate in Arabidopsis: implications for genome size evolution. Genome Biol Evol **5:** 723–733

Yates JR, Ruse CI, Nakorchevsky A (2009) Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng **11:** 49–79

Ye Y, Godzik A (2003) Flexible structure alignment by chaining aligned fragment pairs allowing twists. Bioinformatics (Suppl 2) **19:** ii246–ii255

Yu HN, Wang L, Sun B, Gao S, Cheng AX, Lou HX (2015) Functional characterization of a chalcone synthase from the liverwort Plagiochasma appendiculatum. Plant Cell Rep **34:** 233–245

Yu O, Jez JM (2008) Nature's assembly line: biosynthesis of simple phenylpropanoids and polyketides. Plant J **54:** 750–762

Zhang J, Wang X, Yu O, Tang J, Gu X, Wan X, Fang C (2011) Metabolic profiling of strawberry (Fragaria × ananassa Duch.) during fruit development and maturation. J Exp Bot **62:** 1103–1118