Molecular docking study of *CHS* gene from a medicinally important orchid, *Coelogyne ovalis* Lindl.

Nutan Singh and Suman Kumaria*

Department of Botany, Plant Biotechnology Laboratory, North-Eastern Hill University, Shillong 793 022, India

Chalcone synthase (CHS) is an important enzyme belonging to the polyketide synthase (type III) family. It is well known for substrate specificity and catalyses diverse groups of polyketides of pharmaceutical value. In the present study, the structure of the CHS protein has been predicted from the medicinal orchid, *Coelogyne ovalis* Lindl. This is an evergreen orchid, well known for its medicinal uses. The homology-based model was constructed from CHS of *C. ovalis* (CoCHS) and 16 different ligands were used based on the specificity for molecular docking studies. The four best ligands on the basis of greater negative binding energy were found to be 3-carbomoyl-picoliniyl CoA, followed by carbomoyl-2-napthoyl CoA, *p*-coumaroyl CoA and malonyl CoA. The present structural study reveals that CoCHS has signature and catalytic amino acid residues, namely Cys165, Asn337, His 306, Phe216 and Phe 266. These residues are known to have a broad-range substrate profile and are responsible for the binding of protein ligands. The present study elucidates the chemical basis of CHS from *C. ovalis* by understanding the structural and functional relationship. This provides an insight for manipulating the enzymes, CHS-like for the synthesis of new bioactive compounds, which may further enhance the diversity of the polyketide biosynthetic family.

Keywords: *Coelogyne ovalis*, chalcone synthase, ligands, molecular docking, polyketides.

CHALCONE synthase (CHS; EC 2.3.1.74) is a limiting enzyme of the phenylpropanoid pathway. Two important substrates, viz. malonyl-CoA and coumaryl-CoA in the presence of CHS lead to the production of downstream important bioactive metabolites like flavonoids and anthocyanins. These compounds are important constituents of pharmaceuticals[1,2]. The metabolites in plants are well known for their natural defence mechanisms and are intermediates in the production of many other important products[3]. Due to the diverse structure and various biological as well as pharmacological properties of secondary metabolites, focus has been on the key stimulator enzyme CHS, for enhancement and activation of the secondary metabolites. CHS serves as a starting material for a diverse set of metabolites. These metabolites play a significant role in flower development, fruit pigmentation, UV absorption and as insect repellents and phytoalexins. CHS has been reported to be a gatekeeper enzyme of the phenylpropanoid pathway[4–6].

CHS is known for its wide series of substrate promiscuity and catalytic specificity towards a broad range of thioesters which are involved in the synthesis of a diverse group of medicinally important polyketides[7–11]. It shows a high level of structural similarity with respect to amino acid sequence, and catalytic activities of the enzyme stilbene synthase. The catalytic sites of CHS protein contain conserved amino acid residues, namely cysteine, histidine and asparagine in their active sites. These sites are important for the functionality of this enzyme. Due to its importance, CHS has been extensively studied in several plant species like *Aquilaria sinensis*, *Coelogyne ovalis*, *Oryza sativa*, *Physcomitrella patens*, *Psilotum nudum*, etc.[12–17].

Coelogyne ovalis Lindl., an epiphytic orchid is found dispersed in India as well as in Bhutan, Malaysia, Nepal, Philippines, Sri Lanka and Southwest Pacific Islands[18]. In India, the plants are mostly found in Assam, Meghalaya, Arunachal Pradesh, Sikkim, Kerala and Tamil Naidu. *C. ovalis* is an important medicinal orchid commonly known as 'Jeevanti'. Since ancient times, this plant has been used in traditional medicine system by the indigenous people of South India. The plant parts of *C. ovalis* are regularly used in the treatment of cough, urinary infection and eye disorders[19]. Many bioactive compounds have been reported from *C. ovalis* and these are known for their therapeutic properties. These bioactive compounds include coelogin, flavidin, flavidinin and coeloginin[20,21]. The demand for medicinal plants is increasing worldwide due to the presence of natural bioactive compounds in them, which form an alternative source in modern pharmacological approaches.

In the present study, the 3D structure of CoCHS protein (CHS from *C. ovalis*) has been predicted for molecular docking. The obtained CoCHS protein has been docked with different ligands. This information can be used for the structural alternation of CoCHS protein, and subsequently in the exploration of novel polyketide products.

*For correspondence. (e-mail: sumankhatrikumaria@gmail.com)

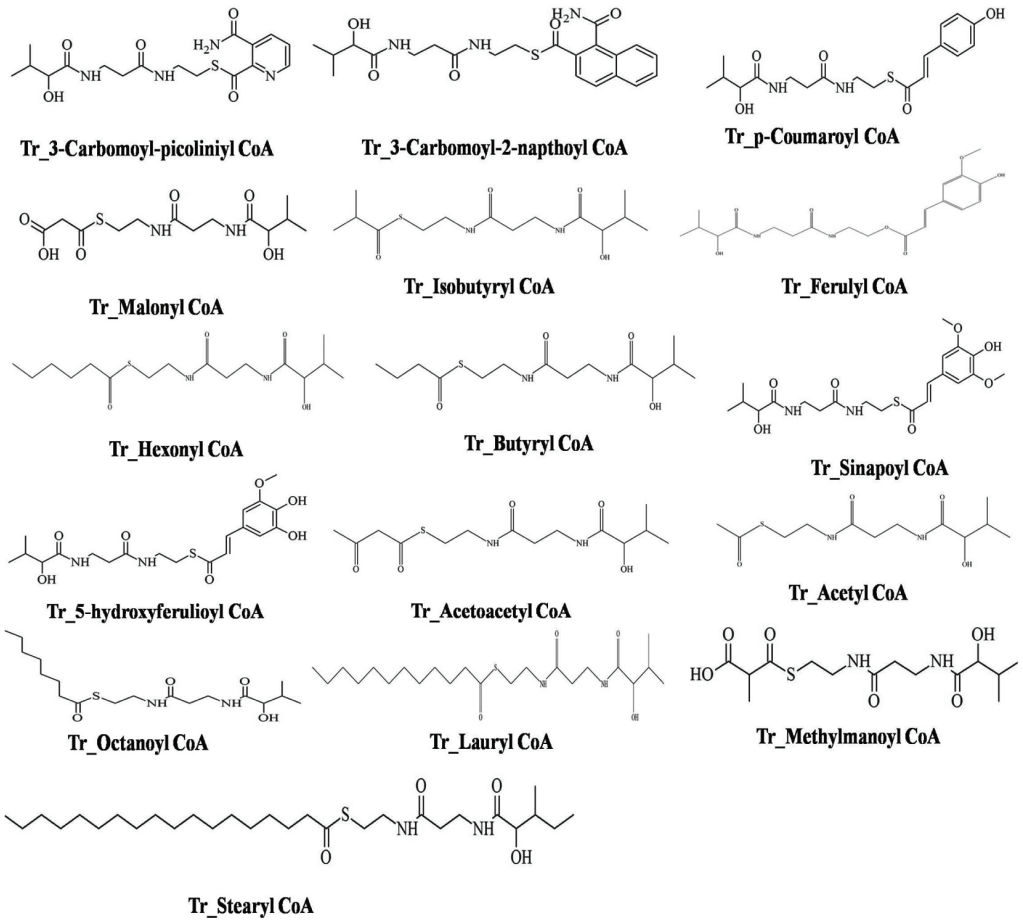

Figure 1. Truncated ligands used in the docking study of *Coelogyne ovalis* chalcone synthase (CoCHS).

Methodology

Sequence alignment and analysis

The isolation and characterization of full-length *CHS* gene sequence from *C. ovalis* has been reported earlier[17]. The obtained nucleotide sequence was verified and deposited at NCBI (accession no. MH796137). The *CoCHS* sequence was translated into the protein sequence and analysed using Expasy (http://www.expasy.org/). The obtained amino acid sequence was BLASTp searched to determine the relevant CHS homology.

Prediction of active sites

Active site prediction server (http://www.scfbio-iitd.res.in/dock/ActiveSite.jsp) was employed to predict the active sites of the macromolecule.

Molecular docking of CoCHS protein

For molecular docking study, homology-based modelling was performed with CoCHS and the template was selected on the basis of high level of sequence identity with that of the CoCHS protein[17]. Subsequently, substrate specificity and ligand interaction with the CoCHS protein were assessed using AutoDock 4.1 (Molecular Graphics Laboratory, CA, USA). After adjusting the parameters of docking, a grid box was drawn around the active site of the enzyme. Sixteen different CoA esters ligands used in this study were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and these were manually further truncated (Figure 1). Ligand-binding plot was constructed using LIGPLOT v.4.5.3 (EMBL-EBL).

Results and discussion

CHS is a well-studied polyketide synthase (type III) enzyme. It is the key rate-limiting enzyme in the pathway which catalyses the branch step and initiates the biosynthesis of flavonoid-based compounds. Flavonoids are essential bioactive metabolites which encompass a diverse collection of valuable compounds, namely chalcones, flavonols, flavones and anthocyanins. These compounds have notable antiviral, antioxidant and other useful biological activities. They also play an important role in the

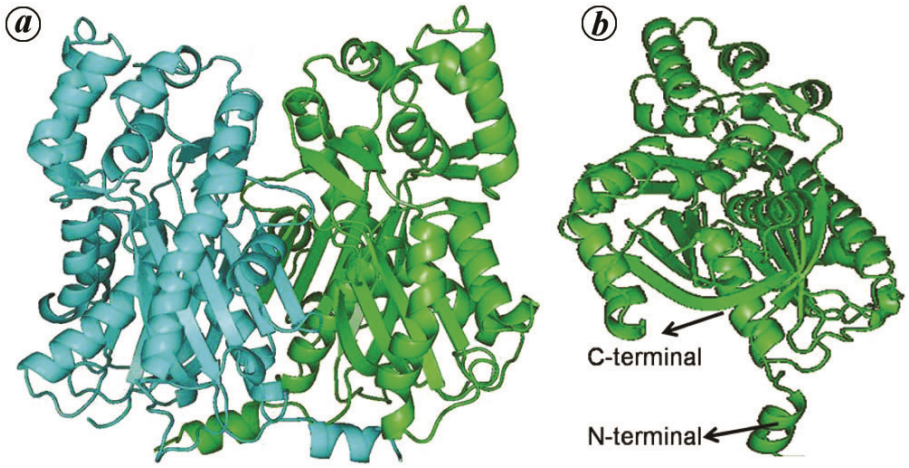

Figure 2. Three-dimensional structure of CoCHS protein: (***a***) homo-dimeric unit and (***b***) monomeric unit.

prevention of cardiovascular diseases, atherosclerotic arteries, etc.[22–24]. Products of the CHS reaction are essential precursors for the production of a large range of novel polyketides which have been reported to be valuable as pharmaceuticals[7,10,25,26]. CHS exhibits inclusive substrate promiscuity and catalytic specificity and catalyses the fabrication of various novel polyketides. Therefore, in this study, *in silico*-based tools were used for understanding the structure of CoCHS protein.

Prediction of three-dimensional structure of CoCHS protein

The three-dimensional structure of CHS protein from *C. ovalis* was predicted using SWISS-MODEL (http://swissmodel.expasy.org/). The oligo state revealed the CoCHS protein to be a homo-dimer consisting of two identical subunits (Figure 2 *a*). Each subunit was found to be made up of 394 amino acid residues (Figure 2 *b*). For homology-based modelling, *Freesia hybrida* CHS protein with PDB ID-4WUM was selected as the template for predicting the structure of CoCHS protein. The 3D structure obtained showed high similarity (82.03%) with the selected template, indicating its high probability of being a CHS protein[17].

Molecular docking analysis of CoCHS

In the present study, 16 different ligands were used to check substrate specificity with the CoCHS protein. Among the ligands tested, 3-carbomoyl-picoliniyl displayed a more negative binding energy (–8.29) compared to the other ligands studied. In 3-carbomoyl-picoliniyl, Asp 306 and Met 60 were the residues involved in the formation of hydrogen bonds with bond length of 2.58 and 2.72 respectively. On the other hand, Phe 216, Phe 266, Cys 165, Asn 332, Gly 212, Asp 208, Leu 268, Val 211, Ala 214, Leu 215, Lys 63, Ser 64, Pro 308, Gly 307, Ala 309 and Asp 306 were the important residues involved in hydrophobic interaction with 3-carbomoyl-picoliniyl. These amino acid residues are reported to be involved in the formation of active sites and might increase the cavity of the binding sites[27]. Subsequently, in the present study successive negative binding energy was reported as –7 for 3-carbomoyl-2-napthoyl CoA, –6.75 for *p*-coumaroyl CoA and –6.25 for malonyl CoA (Table 1). All the ligands were bound to suitable binding pockets in CoCHS protein (Table 1). It is noteworthy to mention that in the present study only the 2D structure of malonyl-CoA was found to allow complete ligand–protein interaction with all the catalytic residues, including the gatekeeper residues of CoCHS protein.

The signature amino acid residues vital to CoCHS protein were Cys165, Asn337, His 306, Phe216 and Phe 266. The other important residues were Glu193, Ile255, Phe216, Phe266, Pro376, Ser134, Ser339, Thr133, Thr195 and Thr198 (refs 27, 28). These residues have been reported to be exceedingly conserved in CHS. The docked conformation of CoCHS protein also contained important active sites, namely Asn337, Cys165 and His304. It has been reported that five amino acid residues, namely Ser134, Glu193, Thr195, Thr198 and Ser339 form the coumaroyl-binding pocket, while, seven amino acids residues, viz. Thr133, Met138, Phe216, Ile255, Gly257, Phe266 and Pro376 form the cyclization pocket[27]. Also, amino acid sequence analysis revealed that it contained all the family signature sequences of CHS functions[17,27,29–31]. The present findings suggest that the isolated CHS sequences of *C. ovalis* are homologous to CHS sequences from other plants.

Figure 1 and Table 1 provide details of the ligands. These ligands bind to the predicted binding pocket of CoCHS macromolecule. The four best docked ligands on the basis of more negative binding energy were 3-carbomoyl-picoliniyl CoA, followed by 3-carbomoyl-2-napthoyl

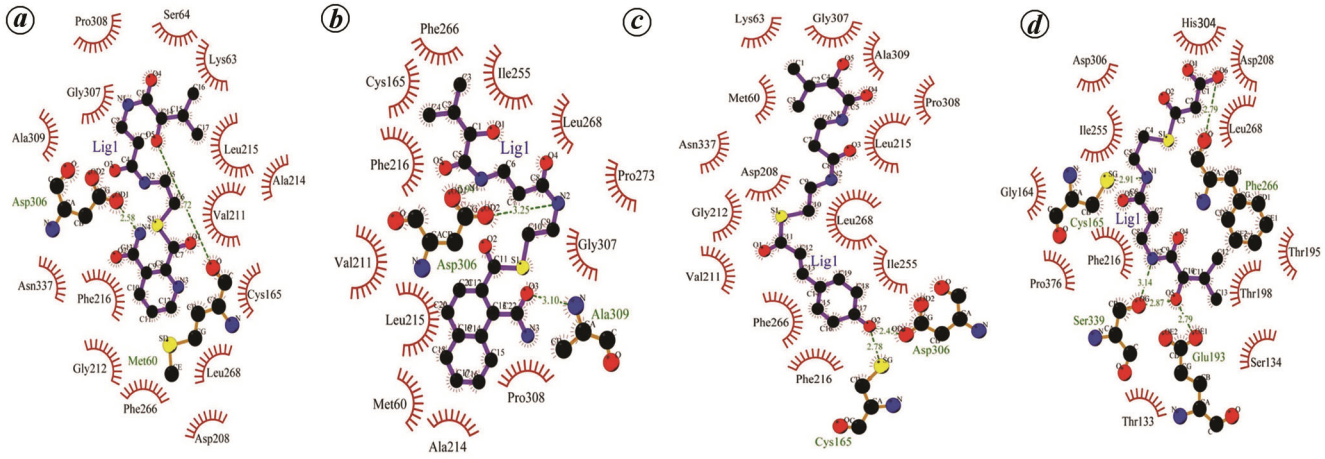

Figure 3. Ligand binding plots: (***a***) Tr_3-carbomoyl-picoliniyl CoA, (***b***) Tr_3-carbomoyl-2-napthoyl CoA, (***c***) Tr_*p*-coumaroyl CoA and (***d***) Tr_malonyl CoA.

Table 1. Docking scores of ligands on modelled structure of *Coelogyne ovalis* chalcone synthase

Ligands	Structure	Docking score
Tr_3-carbomoyl-picoliniyl CoA	CC(C)C(O)C(=O)NCCC(=O)NCCSC(=O)c1ncccc1C(N)=O	–8.29
Tr_3-carbomoyl-2-napthoyl CoA	CC(C)C(O)C(=O)NCCC(=O)NCCSC(=O)c2ccc1ccccc1c2C(N)=O	–7
Tr_*p*-coumaroyl CoA	Oc1ccc(/C=C/C(=O)SCCNC(=O)CCNC(=O)C(O)C(C)C)cc1	–6.75
Tr_malonyl CoA	O=C(CC(=O)O)SCCNC(=O)CCNC(=O)C(O)C(C)C	–6.25
Tr_isobutyryl CoA	O=C(SCCNC(=O)CCNC(=O)C(O)C(C)C)C(C)C	–6.02
Tr_ferulyl CoA	Oc1ccc(cc1OC)/C=C/C(=O)OCCNC(=O)CCNC(=O)C(O)C(C)C	–6.02
Tr_hexonyl CoA	O=C(CCCCC)SCCNC(=O)CCNC(=O)C(O)C(C)C	–5.98
Tr_butyryl CoA	O=C(CCC)SCCNC(=O)CCNC(=O)C(O)C(C)C	–5.91
Tr_sinapoyl CoA	COc1cc(cc(OC)c1O)/C=C/C(=O)SCCNC(=O)CCNC(=O)C(O)C(C)C	–5.84
Tr_5-hydroxyferulioyl CoA	COc1cc(cc(O)c1O)/C=C/C(=O)SCCNC(=O)CCNC(=O)C(O)C(C)C	–5.66
Tr_acetoacetyl CoA	O=C(CC(C)=O)SCCNC(=O)CCNC(=O)C(O)C(C)C	–5.63
Tr_acetyl CoA	O=C(C)SCCNC(=O)CCNC(=O)C(O)C(C)C	–5.16
Tr_octanoyl CoA	O=C(SCCNC(=O)CCNC(=O)C(O)C(C)C)CCCCCCC	–5.11
Tr_lauryl CoA	O=C(CCCCCCCCCCC)SCCNC(=O)CCNC(=O)C(O)C(C)C	–5.01
Tr_methylmanoyl CoA	O=C(SCCNC(=O)CCNC(=O)C(O)C(C)C)C(C)C(=O)O	–4.96
Tr_stearyl CoA	O=C(CCCCCCCCCCCCCCCCC)SCCNC(=O)CCNC(=O)C(O)C(C)C	–3.99

CoA, *p*-coumaroyl CoA and malonyl CoA (Table 1). The highest binding energy score was reported by 3-carbomoyl-picoliniyl CoA (–8.29), followed by carbomoyl-2-napthoyl CoA (–7), *p*-coumaroyl CoA (–6.75) and malonyl CoA (–6.25)[15,29,32,33]. Plots constructed with the four best docked ligands (3-carbomoyl-picoliniyl CoA, 3-carbomoyl-2-napthoyl CoA, *p*-coumaroyl CoA and malonyl CoA using LIGPLOT v.4.5.3 (EMBL-EBL)) showed them binding to the same macromolecule, thus revealing the possible interactions with all the amino acid residues which are important for *CHS* gene functionality (Figure 3).

Conclusion

The present study reveals wide-spectrum substrate specificity of CoCHS protein and the amino acid residues that are crucial to substrate binding. This information can be potentially used for the fabrication of a broad range of phenylpropanoid important polyketide compounds which may be pharmaceutically translated into important drugs for medicinal uses.

1. Koes, R. E., Spelt, C. E. and Mol, J. N., The chalcone synthase multigene family of *Petunia hybrida*: differential, light regulated expression during flower development and UV light induction. *Plant Mol. Biol.*, 1989, **12**, 213–225.
2. Vogt, T., Phenylpropanoid biosynthesis. *Mol. Plant*, 2010, **3**, 2–20.
3. Tohge, T., Yonekura-Sakakibara, K., Niida, R., Wantanabe-Takahasi, A. and Saito, K., Phytochemical genomics in *Arabidopsis thaliana*: a case study for functional identification of flavonoid biosynthesis gene. *Pure Appl. Chem*., 2007, **79**, 811–823.
4. Harborne, J. B. and Turner, B. L. (eds), Application of chemistry at the infraspecific level. In *Plant Chemosystematics*, Academic Press, London, UK, 1984, pp. 237–266.
5. Dixon, R. A., The phytoalexin response: elicitation, signalling and control of host gene expression. *Biol. Rev*., 1986, **61**, 239–291.
6. Schmelzer, E., Jahnen, W. and Hahlbrock, K., *In situ* localization of light-induced chalcone synthase mRNA, chalcone synthase, and flavonoid end products in epidermal cells of parsley leaves. *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 2989–2993.

7. Abe, I., Takahashi, Y. and Noguchi, H., Enzymatic formation of an unnatural C6–C5 aromatic polyketide by plant type III polyketide synthases. *Org. Lett*., 2002, **4**, 3623–3626.
8. Dao, T. T., Linthorst, H. J. and Verpoorte, R., Chalcone synthase and its functions in plant resistance. *Phytochem. Rev*., 2011, **10**, 397–412.
9. Morita, H., Takahashi, Y., Noguchi, H. and Abe, I., Enzymatic formation of unnatural aromatic polyketides by chalcone synthase. *Biochem. Biophys. Res. Commun*., 2000, **279**, 190–195.
10. Oguro, S., Akashi, T., Ayabe, S. I., Noguchi, H. and Abe, I., Probing biosynthesis of plant polyketides with synthetic N-acetylcysteamine thioesters. *Biochem. Biophys. Res. Commun.*, 2004, **325**, 561–567.
11. Reddy, A. R., Scheffler, B., Madhuri, G., Srivastava, M. N., Kumar, A., Sathyanarayanan, P. V. and Mohan, M., Chalcone synthase in rice (*Oryza sativa* L.): detection of the CHS protein in seedlings and molecular mapping of the *CHS* locus. *Plant Mol. Biol*., 1996, **32**, 735–743.
12. Awasthi, P., Mahajan, V., Jamwal, V. L., Kapoor, N., Rasool, S., Bedi, Y. S. and Gandhi, S. G., Cloning and expression analysis of chalcone synthase gene from *Coleus forskohlii. J. Genet*., 2016, **95**, 647–657.
13. Goodwin, P. H., Hsiang, T. and Erickson, L., A comparison of stilbene and chalcone synthases including a new stilbene synthase gene from *Vitis riparia* cv. Gloire de Montpellier. *Plant Sci.*, 2000, **151**, 1–8.
14. Yamazaki, Y. *et al.*, Diverse chalcone synthase superfamily enzymes from the most primitive vascular plant, *Psilotum nudum*. *Planta*, 2001, **214**, 75–84.
15. Jiang, C., Schommer, C. K., Kim, S. Y. and Suh, D. Y., Cloning and characterization of chalcone synthase from the moss, *Physcomitrella patens*. *Phytochemistry*, 2006, **67**, 2531–2540.
16. Wang, M. X., Li, W. L., Zhang, Z., Wei, J. H., Yang, Y., Xu, Y. H. and Liang, L., Cloning and bioinformatics analysis of chalcone synthase (*AsCHS1*) gene in *Aquilaria sinensis*. *China J. Chin. Mater. Med.*, 2013, **38**, 149–153.
17. Singh, N. and Kumaria, S., Molecular cloning and characterization of chalcone synthase gene from *Coelogyne ovalis* Lindl. and its stress-dependent expression. *Gene*, 2007, **762**, 145104.
18. Nongrum, I., Kumaria, S. and Tandon, P., Influence of *in vitro* media on asymbiotic germination, plantlet development and *ex vitro* establishment of *Coelogyne ovalis* and *Coelogyne nitida*. *Proc. Indian Natl. Sci. Acad*., 2007, **73**, 205–207.
19. Yonzone, R., Lama, D., Bhujel, R. B. and Rai, S., Present availability status, diversity resources and distribution of medicinal orchid species in Darjeeling Himalaya of West Bengal, India. *Int. J. Pharm. Nat. Med*., 2013, **1**, 14–35.
20. Majumder, P. L. and Laha, S., Occurrence of 2,7-dihydroxy-3,4-6-trimethoxy-9,10-dihydropenanthrene in *Coelogyne ovalis*, a high altitude Himalayan orchid: application of C-13 NMR spectroscopy in structure elucidation. *J. Indian Chem. Soc.*, 1981, **58**, 928–929.
21. Sachdev, K. and Kulshreshtha, D. K., Phenolic constituents of *Coelogyne ovalis*. *Phytochemistry*, 1986, **25**, 499–502.
22. Commenges, D., Scotet, V., Renaud, S., Jacqmin-Gadda, H., Barberger-Gateau, P. and Dartigues, J. F., Intake of flavonoids and risk of dementia. *Eur. J. Epidemiol*., 2000, **16**, 357–363.
23. Knekt, P., Kumpulainen, J., Järvinen, R., Rissanen, H., Heliövaara, M., Reunanen, A. and Aromaa, A., Flavonoid intake and risk of chronic diseases. *Am. J. Clin. Nutr.*, 2002, **76**, 560–568.
24. Teixeira, S., Siquet, C., Alves, C., Boal, I., Marques, M. P., Borges, F. and Reis, S., Structure–property studies on the antioxidant activity of flavonoids present in diet. *Free Radic. Biol. Med.*, 2005, **39**, 1099–1108.
25. Abe, I., Watanabe, T., Morita, H., Kohno, T. and Noguchi, H., Engineered biosynthesis of plant polyketides: manipulation of chalcone synthase. *Org. Lett*., 2006, **8**, 499–502.
26. Jez, J. M., Bowman, M. E. and Noel, J. P., Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 5319–5324.
27. Ferrer, J. L., Jez, J. M., Bowman, M. E., Dixon, R. A. and Noel, J. P., Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nature Struct. Mol. Biol*., 1999, **6**, 775.
28. Ma, W., Wu, Y., Wu, M., Ren, Z. and Zhong, Y., Cloning, characterization and expression of chalcone synthase from medicinal plant *Rhus chinensis*. *J. Plant Biochem. Biotechnol*., 2015, **24**, 18–24.
29. Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R. J., Schröder, J. and Teeri, T. H., Chalcone synthase-like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (Asteraceae). *Plant Mol. Biol*., 1995, **28**, 47–60.
30. Schröder, J., A family of plant-specific polyketide synthases: facts and predictions. *Trends Plant Sci*., 1997, **2**, 373–378.
31. Kim, S. H., Mizuno, K. and Fujimura, T., Regulated expression of ADP glucose pyrophosphorylase and chalcone synthase during root development in sweet potato. *Plant Growth Regul*., 2002, **38**, 173–179.
32. Han, Y. Y., Ming, F., Wang, J. W., Wen, J. G., Ye, M. M. and Shen, D. L., Cloning and characterization of a novel chalcone synthase gene from *Phalaenopsis* hybrida orchid flowers. *Russ. J. Plant Physiol*., 2006, **53**, 223–230.
33. Wannapinpong, S., Srikulnath, K., Thongpan, A., Choowongkomon, K. and Peyachoknagul, S., Molecular cloning and characterization of the CHS gene family in turmeric (*Curcuma longa* Linn.). *J. Plant Biochem. Biotechnol*., 2015, **24**, 25–33.

ACKNOWLEDGEMENTS. We acknowledge financial support from the Department of Biotechnology, Government of India (BT/PR16385/NER/95/124/2015); University Grants Commission, Non-NET fellowship from North-Eastern Hill University (NEHU), Shillong (F24-25/DSW/2014/864). We thank Dr Atanu Bhattacharjee and Dr Arun Gurung (NEHU, Shillong) for help during the experiments.

Received 21 May 2021; accepted 23 August 2021

doi: 10.18520/cs/v121/i10/1323-1327