Homology modeling and docking study of chalcone synthase gene (CfCHS) from Coleus forskohlii

Praveen Awasthi¹,²*, Vijay Lakshmi Jamwal¹, Nitika Kapoor¹ and Shafaq Rasool²

Abstract: Flavonoids are important secondary metabolites in plants. Chalcone synthase (CHS) catalyzes the first committed step in the flavonoids biosynthesis. CHS belongs to type III polyketide synthases that are known for their broad substrate specificity and catalytic potential toward a wide range of thioesters, to produce diverse novel polyketides of pharmaceutical importance. In this study, an in silico approach was used to understand the structure and function of CHS protein from an important medicinal plant Coleus forskohlii. A homology model of CfCHS was built and docking studies were carried out using 25 ligands. Best four docked ligands in proposed binding pocket of CfCHS were: Cinnamoyl CoA, 2-Carbamoylbenzoyl CoA, Benzoyl CoA and p-Coumaroyl CoA. Cys 164, His 304, and Asn 337 were found to be catalytic residues of CfCHS. Further two important residues, Phe 216 and Phe 266, were found to be the gatekeeper residues involved in π–π interaction with ligands. Present study revealed broad spectrum substrate profile of CfCHS and important key residues involved in substrate binding. This is the first report of homology modeling and docking analysis of CHS from C. forskohlii.

Subjects: Bioinformatics; Biology; Biotechnology

Keywords: CHS; flavonoid; polyketide synthase

ABOUT THE AUTHORS

Praveen Awasthi, is a research scholar at CSIR-Indian Institute of Integrative Medicine, Jammu (India). His area of research is plant molecular biology with a keen interest in the study of secondary metabolic pathways in medicinal plants. He has been involved in various research projects on important medicinal plants such as Coleus forskohlii, Ocimum basilicum, Zingiber officinale, etc. He has also worked on understanding the microsatellite dynamics within Brassicaceae family.

Nitika Kapoor and Vijay Lakshmi Jamwal are research scholars at the CSIR-Indian Institute of Integrative Medicine, Jammu (India). They are working on understanding various aspects of secondary metabolism in medicinal and aromatic plants.

Shafaq Rasool is an assistant professor at Shri Mata Vaishno Devi University, India. Her areas of specialization are molecular biology, microbial biotechnology, and protein engineering.

PUBLIC INTEREST STATEMENT

The functional diversity of the plant-specific type III polyketide synthases (PKSs) is remarkable. Chalcone synthase (CHS) is one of the important class of type III PKS, that show extremely broad substrate promiscuity and catalytic potential toward a wide range of substrates from aromatic to aliphatic thioesters, to produce diverse novel polyketides of pharmaceutical importance. In this study, we attempted to understand the substrate diversity of CHS from an important medicinal plant—C. forskohlii.
1. Introduction

Flavonoids are ubiquitous in higher plants and play significant roles in a wide array of processes like pollination, floral pigmentation, nitrogen fixation, and are produced in response to stress, UV radiation, pathogens, insects, etc. (Fini, Brunetti, Di Ferdinando, Ferrini, & Tattini, 2011; Lillo, Lea, & Ruoff, 2008; Pásold, Siegel, Seidel, & Ludwig-Müller, 2010; Taylor & Grotewold, 2005; Winkel-Shirley, 2002). Flavonoids are known to have a close connection with human health. Besides the remarkable antioxidant, antiviral, and anti-proliferative biological activities, they have also been used for treatment of cardiovascular disease (Commenges et al., 2000; Knekt et al., 2002; Teixeira et al., 2005).

CHS, the well-studied plants-specific type III PKS, is an important enzyme that channels the flux of phenylpropanoid pathway toward biosynthesis of flavonoids (Schröder, 1997). It catalyzes the condensation of three molecules of malonyl-CoA and one molecule of p-coumaroyl-CoA, yielding naringenin chalcone (Martens & Mithöfer, 2005). CHS shows broad substrate promiscuity and catalytic potential toward a wide range of thioesters (benzoyl-CoA, feruloyl-CoA, phenyl acetyl-CoA, hexanoyl-CoA, and isovaleryl-CoA and many more) to produce diverse novel polyketides of pharmaceutical importance (Abe, Takahashi, & Noguchi, 2002; Abe, Watanabe, & Noguchi, 2004; Dao, Linthorst, & Verpoorte, 2011; Morita, Takahashi, Noguchi, & Abe, 2000; Oguro, Akashi, Ayabe, Noguchi, & Abe, 2004). CHS proteins share high degree of similarity in their amino acid sequence, structure, and general catalytic principles, including the presence of the conserved Cys–His–Asn catalytic triad in the buried active site. CHS has been studied in many plant species including Oryza sativa L. (Poaceae), Aquilaria sinensis (Lour) Spreng. (Thymelaeaceae), Physcomitrella patens (Hedw.) Bruch & Schimp. (Funariaceae), Psilotum nudum (L.) P. Beauv. (Psilotaceae), etc. (Goodwin, Hsiang, & Erickson, 2000; Jiang, Schommer, Kim, & Suh, 2006; Reddy et al., 1996; Wang et al., 2013; Yamazaki et al., 2001). Coleus forskohlii (Willd.) Briq. (Lamiaceae) is an important medicinal plant in Indian system of medicine (Kavitha, Rajamani, & Vadivel, 2010). The plant is being used from ancient time in ayurvedic traditional medicine for treatment of diverse human health disorders (Lukhoba, Simmonds, & Paton, 2006). Not many flavonoids have been reported from C. forskohlii (Alasbahi & Melzig, 2010), so it would be interesting to analyze the substrate diversity of CfCHS, as this may help in providing clues about the possible products that may be formed by CfCHS in planta. This information could also be exploited for in vitro biotechnological use of CfCHS for production of novel polyketide products.

2. Methodology

2.1. Amino acid sequence analysis, template searching, and sequence alignment

In our previous study, we have cloned the full-length gene sequence of CfCHS gene (Awasthi et al., 2016) and submitted the sequence to NCBI database (accession no. KF643243). CfCHS sequence was analyzed using bioinformatics tools at the website http://www.expasy.org/. The protein sequence of CfCHS was analyzed using BLASTP tool to identify closely related CHS homologs. The sequence that showed the highest sequence identity with CfCHS was selected as a template for homology modeling of CfCHS protein.

2.2. Homology modeling and structure analysis of CfCHS protein

The Protein model of CfCHS was prepared using homology modeling module of Schrödinger Suite (Maestro, version 9.4, Schrödinger, LLC, New York, NY). The 3D structure of the protein was predicted using crystal structure of CHS (PDB ID: 4WUM) as a template. To check the stereochemical stability of protein model, ramachandran plot was obtained from procheck module of the SAVES server (http://services.mbi.ucla.edu/SAVES/).

2.3. Molecular docking

CfCHS protein model was prepared for docking using protein preparation wizard of Schrödinger Suite (Maestro, version 9.4, Schrödinger, LLC, New York, NY). The active site of protein model was predicted using active site prediction server (http://www.scfbio-iitd.res.in/dock/ActiveSite.jsp) and the grid was generated around it. Glide XP precision mode (Glide, version 5.9, Schrödinger, LLC, New York, NY) was
used to carry out the molecular docking of CfCHS protein (Abe & Morita, 2010; Friesner et al., 2006; Halgren et al., 2004; Sastry, Adzhigirey, Day, Annabhimoju, & Sherman, 2013). A total of 25 ligands (CoA esters) were downloaded from the pubchem database (https://pubchem.ncbi.nlm.nih.gov/) and were manually truncated using Schrondinger suite 2013. The structures of these ligands are given in Figure 1.

To validate the methodology used here for study of CFCHS, we have modeled four functionally characterized type III PKSs that are; chalcone synthase (Gb_CHS) from Gerbera hybrida (Mutisieae), stilbene synthase (Pd_STS) from Pinus densiflora Siebold & Zucc. (Pinaceae), 2-pyrone synthase (Gb_PS) from G. hybrida and valerophenone synthase (HL_VS) from Humulus lupulus L. (Cannabaceae). Truncated CoA esters were taken as ligand (Tr-Acetyl CoA, Tr- p-Coumaroyl CoA, Tr-p-Isovaleroyl CoA). Molecular modeling and docking experiment was performed as discussed above.

Figure 1. Ligands of CFCHS. Structure of 25 ligands used in docking study of CFCHS in this study.
3. Results

3.1. Homology modeling and stereochemical analysis of CfCHS

Homology modeling for CfCHS was carried out using CHS (PDB ID: 4WUM) from Freesia hybrida as a template. Protein model was developed using Schrödinger Suite (Maestro, version 9.4, Schrödinger, LLC, New York, NY) (Figure 2(a)). Stereochemical analysis of the predicted model of CfCHS protein was analyzed by procheck server (http://services.mbi.ucla.edu/SAVES/). Ramachandran plot analysis of CfCHS showed 90.3% residues in the most favorable region, 8.2% residues in the additional allowed region, 1.5% in the generously allowed region, and 0.0% in the disallowed region (Figure 2(b)). The results of the procheck analysis indicated that no residue had phi/psi angles in the disallowed regions suggesting the acceptability of ramachandran plot for CfCHS protein (Figure 2(b)). For good protein model, 90% of residues should be in the most favorable region (Morris, MacArthur, Hutchinson, & Thornton, 1992). 91.75% of residues of CfCHS had an average 3D-1D score $\geq$ 0.2. The protein structure was further analyzed for Z-score mean, Z-score Stdev, and Z-score RMS (Supplementary Figure 1).

3.2. Molecular docking study of CfCHS

Firstly, methodology used for docking experiment was validated through study of some of the known type III PKS. As we know, CHSs and STSs (stilbene synthase) prefer p-coumaroyl-CoA, 2-pyrone synthase from G. hybrida prefers acetyl-CoA, and valerophenone synthase from H. lupulus prefers isovaleryl-CoA. Our docking analysis concurs with the experimentally characterized enzymatic activities of these proteins, thus validating our methodology (Supplementary Table 1, Supplementary Figure 2). This methodology was previously also employed to predict substrate selectivity of similar enzymes (Bhat et al., 2013).

Docking experiments were performed with the CfCHS protein model using CoA-esters of cinnamic acid and its derivatives as ligand data-set (Figure 1). Ligands that showed high binding affinity for CfCHS (Table 1; Ligands: Tr_Cinnamoyl CoA, Tr_2-Carbamoylbenzoyl CoA, Tr_Benzoyl CoA, and Tr_p-Coumaroyl CoA) were also found to be suitably placed in the proposed binding pocket of CfCHS (Figure 3).

It is proposed that Phe 266 is an important residue that interacts with Tr_Cinnamoyl CoA (Figure 4). Tr_Cinnamoyl CoA was found to have $\pi-\pi$ interaction with Phe 266 ($\sim$5.31 Å), while Lys 270 (1.88 Å) was involved in hydrogen bond formation with ligand Tr_cinnamoyl CoA.
The docking score for Tr_2-Carbamoylbenzoyl CoA was −8.578 (Table 1). Residue which was involved in substrate (Tr_2-Carbamoylbenzoyl CoA) binding through hydrogen bonds was Asn 337 with bond lengths 2.18 and 2.49 Å (Figure 4).

Tr_Benzoyl CoA was found to be the one of the preferred substrate for CfCHS among all 25 docked molecules, having a docking score of −8.487 (Table 1). Phe 266 was found to be an important residue of CfCHS protein that interacted with Tr_Benzoyl CoA. Asn 337 (2.16 Å) was found to be involved in hydrogen bond formation while Phe 266 (5.44 Å) had π–π interaction with the ligand (Figure 4).

p-Coumaroyl CoA is one of the most common substrate for CHS in plants (Dao, Linthorst, & Verpoorte, 2011). Tr_p-Coumaroyl CoA was found to be best docked in the proposed binding pocket of CfCHS in this study. Important catalytic residue Cys 164 was found to be involved in hydrogen bond formation with bond length of 2.47 Å (Figure 4).

4. Discussion
The functional diversity of the plant-specific type III PKSs is remarkable. CHS is an important class of type III PKS, which shows extremely broad substrate specificity and catalytic potential toward a wide range of substrates, from aromatic to aliphatic thioesters, to catalyze the formation of diverse and novel polyketides of pharmaceutical importance (Abe et al., 2002, 2004; Abe, Noma, Noguchi, & Abe, 2002, 2004; Abe, Noma, Noguchi, & Abe, 2002, 2004).
In this study, we attempted to understand the substrate profile of CHS from an important medicinal plant, *C. forskohlii*. This is a first report of modeling and docking analysis of CfCHS from this herb.Docking results for 25 probable substrates for CHS reveal the broad spectrum catalytic potential of CfCHS toward different ligands. In order to find out the key residues involved in substrate interaction, we selected four best docked molecules (Tr_Cinnamoyl CoA, Tr_2-Carbamoylbenzoyl CoA, Tr_Benzoyl CoA, and Tr_p-Coumaroyl CoA). Cys 164, His 304, and Asn 337 were found to be present among the residues interacting with substrates under study. As per previous studies with CHS from other plants, Cys, His, and Asn were identified as key residues in the catalytic triad present in the CHS (Abe & Morita, 2010). Apart from this, two other residues (Phe 266 and Phe 216) were found to have interaction with docked molecules. These are known as gatekeeper residues of CHS (Jez & Noel, 2000). In CfCHS, Phe 266 was found to have strong π–π interaction with all four best docked molecules. In this study, apart from above five important residues, Lys 270 were also found to be involved in hydrogen bond formation. Lys 270 was involved in hydrogen bond formation with substrate Tr_Cinnamoyl CoA.

This study disclosed the broad spectrum substrate profile of CfCHS and important key residues involved in substrate binding. This knowledge can be used for the production of a wide range of pharmaceutically important polyketides.

![Molecular docking study of CfCHS. Diagram showing (a) Tr_Cinnamoyl CoA, (b) Tr_2-Carbamoylbenzoyl CoA, (c) Tr_Benzoyl CoA, and (d) Tr_p-Coumaroyl CoA docked into the proposed binding pockets of CfCHS.](image-url)
Figure 4. 2D representation of ligand interaction of CfCHS. 2D representation of ligand interaction of CfCHS; (a) Tr_Cinnamoyl CoA, (b) 2-Carbamoylbenzoyl CoA, (c) Tr_Benzoyl CoA, and (d) Tr_p-Coumaroyl CoA. Arrow indicates the interaction between ligands and respective amino acids.

Supplementary material
Supplementary material for this article can be accessed here http://dx.doi.org/10.1080/23312025.2016.1175332.

Acknowledgments
Authors are thankful to Director, Indian Institute of Integrative Medicine (CSIR), for providing the facilities for carrying out the work. Authors thank Ms Vidushi Mahajan, who carried out full-length sequencing of CfCHS gene. Authors express sincere gratitude to Dr Sumit G Gandhi for his guidance and for correcting the manuscript and figures. Authors also thank Dr Y. S Bedi for his support and help in the study.

Funding
This work was supported by CSIR.

Competing interests
The authors declare no competing interests.

Author details
Praveen Awasthi1,2 E-mail: praveenawasthi1984@gmail.com
Vijay Lakshmi Jamwal1 E-mail: vijaylakshmjamwal@gmail.com
Nitika Kapoor1 E-mail: k.niti21@gmail.com
Shafaq Rasool2 E-mail: shafaq.rasool@smvdu.ac.in

1 Plant Biotechnology Division, Indian Institute of Integrative Medicine (CSIR-IIIM), Council of Scientific and Industrial Research, Canal Road, Jammu 180001, India.
2 Department of Biotechnology, Shri Mata Vaishno Devi University, Katra 182301, India.

Author contributions
Praveen Awasthi carried out the study, prepared the figures, and wrote the manuscript. Vijay Lakshmi Jamwal and Nitika Kapoor helped Praveen Awasthi in carrying out the analysis. Shafaq Rasool carried out the proofread and provided valuable input in the manuscript.

Citation information
Cite this article as: Homology modeling and docking study of chalcone synthase gene (CfCHS) from Coleus forskohlii, Praveen Awasthi, Vijay Lakshmi Jamwal, Nitika Kapoor & Shafaq Rasool, Cogent Biology (2016), 2: 1175332.

References
Abe, I., & Morita, H. (2010). Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. Natural Product Reports, 27, 809–838. http://dx.doi.org/10.1039/b009988n
Abe, I., Tokashiki, Y., & Noguchi, H. (2002). Enzymatic formation of an unnatural C6−C5 aromatic polyketide by plant type III polyketide synthases. Organic Letters, 4, 3623–3626. http://dx.doi.org/10.1021/ol0201409
Abe, I., Watanabe, T., & Noguchi, H. (2004). Enzymatic formation of long-chain polyketide pyrones by plant type III polyketide synthases. Phytochemistry, 65, 2447–2453. http://dx.doi.org/10.1016/j.phytochem.2004.08.005

Abe, T., Noma, H., Noguchi, H., & Abe, I. (2006). Enzymatic formation of an unnatural methylated triketide by plant type III polyketide synthases. Tetrahedron Letters, 47, 8727–8730. http://dx.doi.org/10.1016/j.tetlet.2006.10.012

Alasbahi, R. H., & Melzig, M. F. (2010). Plectranthus barbatus: A review of phytochemistry, ethnobotanical uses and pharmacology—Part I. Planta Medica, 76, 653–661. http://dx.doi.org/10.1055/s-0030-1240898

Awasthi, P., Mahajan, V., Jamwal, V. L., Kapoor, N., Bedi, Y. S., & Gandhi, S. G. (2016). Cloning and expression analysis of chalcone synthase gene from Coleus forskohlii. Journal of Genetics. Retrieved from http://www.ias.ac.in/public/Resources/General/jogen/jgen-15-498-ue.pdf

Bhat, W. W., Dhar, N., Razdan, S., Rani, S., Mehra, R., Nargotra, A., … Lattoo, S. K. (2013). Molecular characterization of UGT19F4 and UGT86C4, two glycosyltransferases from Picrohiza kurrooa: Comparative structural insight and evaluation of substrate recognition. PlaO NE, 8, e73804. http://dx.doi.org/10.1371/journal.pone.0073804

Commenges, D., Scotet, V., Renaud, S., Jacquim-Gadda, H., Barberger-Gateau, P., & Dartigues, J. F. (2000). Intake of flavonoids and risk of dementia. European Journal of Epidemiology, 16, 357–363. http://dx.doi.org/10.1023/A:1002841613711

Dao, T. T. H., Linthorst, H. J. M., & Verpoorte, R. (2011). Type III polyketide synthases. Phytochemistry, 66, 2399–2407. http://dx.doi.org/10.1016/j.phytochem.2005.07.013

Morita, H., Takahashi, Y., Noguchi, H., & Abe, I. (2000). Enzymatic formation of unnatural aromatic polyketides by chalcone synthase. Biochemical and Biophysical Research Communications, 279, 190–195. http://dx.doi.org/10.1006.bbrc.1999.0074

Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. Proteins: Structure, Function, and Genetics, 12, 345–366. http://dx.doi.org/10.1006/psfg.1992.1003

Oguro, S., Akashi, T., Ayate, S.-I., Noguchi, H., & Abe, I. (2004). Probing biosynthesis of plant polyketides with synthetic N-acetylcysteamine thioesters. Biochemical and Biophysical Research Communications, 325, 561–567. http://dx.doi.org/10.1016/j.bbrc.2004.10.057

Pittock, S., Siegel, I., Seidel, C., & Ludwig-Müller, J. (2010). Flavonoid accumulation in Arabidopsis thaliana root galls caused by the obligate biotrophic pathogen Phasmosphora brassicae. Molecular Plant Pathology, 11, 545–562. http://dx.doi.org/10.1111/j.1364-3703.2010.00628.x

Reddy, A. R., Scheffler, B., Madhuri, G., Srivastava, M. N., Kumar, A., Sathyarayanan, P. V., Mohan, M. (1996). Chalcone synthase in rice (Oryza sativa L.): Detection of the CHS protein in seedlings and molecular mapping of the chs locus. Plant Molecular Biology, 32, 735–743. http://dx.doi.org/10.1007/BF00002014

Sosstry, G. M., Adzhigirey, M., Day, T., Annabihu, M., & Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. Journal of Computer-Aided Molecular Design, 27, 221–234. http://dx.doi.org/10.1007/s10822-013-9644-8

Schödter, J. (1997). A family of plant-specific polyketide synthases: Facts and predictions. Trends in Plant Science, 2, 373–378. http://dx.doi.org/10.1016/S1360-1385(97)87121-X

Taylor, L. P., & Groteveld, E. (2005). Flavonoids as developmental regulators. Current Opinion in Plant Biology, 8, 317–323. http://dx.doi.org/10.1016/j.pbi.2005.03.005

Teixeira, S., Siquet, C., Alves, C., Boal, I., Marques, M. P., Borges, F., … Reis, S. (2005). Structure-property studies on the antioxidant activity of flavonoids present in diet. Free Radical Biology & Medicine, 39, 1099–1108. http://dx.doi.org/10.1016/j.freeradbiomed.2005.08.004

Wang, M.-X., Li, W.-L., Zhang, Z., Wei, J.-H., Yang, Y., Xu, Y.-H., … Sankawa, U. (2001). Diverse chalcone synthase (AsCHS1) gene in Aquilaria sinensis. Phytochemistry, 58, 741–749. http://dx.doi.org/10.1016/s0031-9422(00)00217-7

Yamazaki, Y., Suh, D.-Y., Sitthithaworn, W., Ishiguro, K., Kobayashi, Y., Shibuya, M., … Sankawa, U. (2001). Diverse chalcone synthase superfamily enzymes from the most primitive vascular plant, Pilsotium nudum. Planta, 214, 75–84.
