Engineering a plant polyketide synthase for the biosynthesis of methylated flavonoids

Bo Peng[a], Lili Zhang[b], Siqi He[a], Rick Oerlemans[b], Wim J. Quax[a], Matthew R. Groves[b], Kristina Haslinger[b]

[a]Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713AV Groningen, The Netherlands

[b]XB20 Drug Design, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713AV Groningen, The Netherlands

Correspondence: k.haslinger@rug.nl

Abstract

Homoeriodictyol and hesperetin are naturally occurring O-methylated flavonoids with many health-promoting properties. They are produced in plants in low abundance and as complex mixtures of similar compounds that are difficult to separate. Synthetic biology offers the opportunity to produce various flavonoids in a targeted, bottom-up approach in engineered microbes with high product titers. However, the production of O-methylated flavonoids is currently still highly inefficient.

In this study, we investigated and engineered a combination of enzymes that had previously been shown to support homoeriodictyol and hesperetin production in *Escherichia coli* from fed O-methylated cinnamic acids. We determined the crystal structures of the enzyme catalyzing the first committed step of the pathway, chalcone synthase from *Hordeum vulgare*, in three ligand-bound states. Based on these structures and a multiple sequence alignment with other chalcone synthases, we constructed mutant variants and assessed their performance in *E. coli* towards producing methylated flavonoids. With our best mutant variant, HvCHS (Q232P, D234V), we were able to produce homoeriodictyol and hesperetin at 2 times and 10 times higher titers than previously reported. Our findings will facilitate the further engineering of this enzyme towards higher production of methylated flavonoids.

Keywords

enzyme engineering, crystal structure, homoeriodictyol, hesperetin, chalcone synthase, *Escherichia coli*

30 ## 1. Introduction

31 Flavonoids are natural polyphenolic compounds ubiquitously found in various flowers, fruits,
32 and vegetables[1,2]. In nature, flavonoids are important for plant growth and reproduction, for
33 attracting pollinators, and for protecting against biotic and abiotic stresses, such as harmful
34 ultraviolet radiation[3,4]. More than 15,000 compounds have been identified to date and many
35 have been shown to have health-promoting effects[5], such as anti-cancer, anti-inflammatory,
36 anti-mutagenic, and anti-oxidative properties[6]. These health-promoting effects make
37 flavonoids an attractive ingredient for nutraceutical, pharmaceutical, and cosmetic
38 applications[1].

39 Homoeriodictyol and hesperetin are important, naturally occurring O-methylated flavonoids,
40 which exhibit higher biological activities and better pharmacological properties, including
41 metabolic stability, membrane transport capability, and oral bioavailability, compared to
42 unmethylated flavonoids[7]. Homoeriodictyol, which is O-methylated in the 3'-position, was
43 previously isolated from *Viscum articulatum* Burm and has been shown to have anti-cancer
44 effects in MCF-7, HeLa, and HT-29 cells[8]. Furthermore, it was shown to protect endothelial
45 cells from oxidative stress and mitochondrial dysfunction[9]. Hesperetin, which is O-methylated
46 in the 4'-position, has been reported to ameliorate Alzheimer's disease[10], and inhibit the
47 migration of breast cancer[11]. More importantly, hesperetin displays strong inhibitory activity
48 towards several viruses, such as influenza A 14 virus, parainfluenza virus type-3, and SARS-
49 CoV[12,13].

50 The current industrial production of flavonoids mainly relies on extraction from plants. This
51 strategy has the limitations that the flavonoid content and yield is variable between seasons,
52 and that flavonoids exist in complex mixtures that need to be separated with large volumes
53 of organic solvents. The chemical synthesis of flavonoids requires harsh reaction conditions
54 and toxic substrates, which are not eco-friendly[14].

55 Another promising route for generating flavonoids is the biosynthesis of flavonoids in

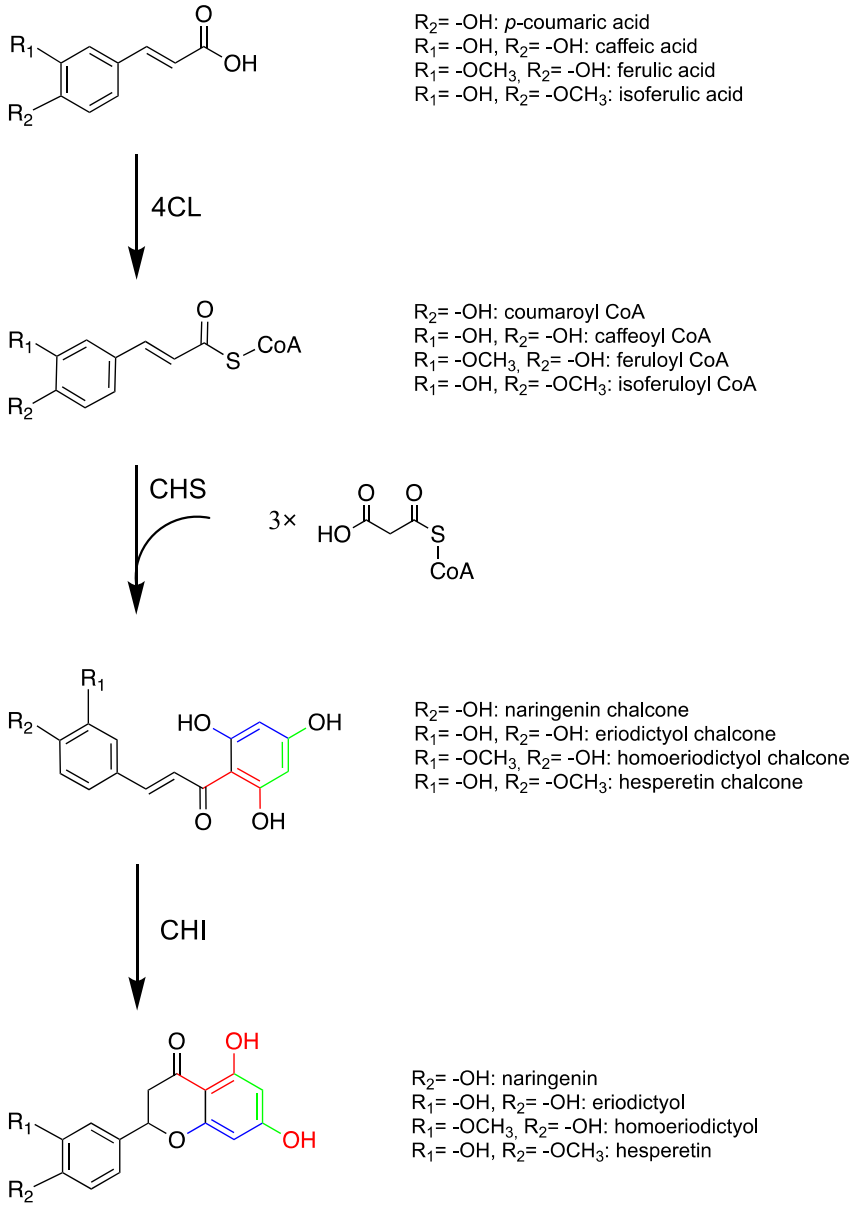
56 microbial cell factories. The rapid development of synthetic biology has enabled the
57 heterologous expression of the flavonoid biosynthesis pathway in microorganisms such as
58 *E. coli* and *Saccharomyces cerevisiae*[15–18]. In the first step, *p*-coumaric acid is activated by 4-
59 coumarate-coenzyme A (CoA) ligase (4CL) to give coumaroyl-CoA (Fig. 1). Then, chalcone
60 synthase (CHS) condenses one coumaroyl-CoA and three malonyl-CoA to form naringenin
61 chalcone, which is subsequently converted into the (2S)-naringenin by chalcone isomerase
62 (CHI). In plants, naringenin can be converted into other flavonoids by tailoring enzymes and is
63 thus a key intermediate in the pathway. CHS is the first committed enzyme in flavonoid
64 synthesis and belongs to the superfamily of type three polyketide synthases[19]. It has been
65 reported that altering certain residues of such plant type three polyketide synthases improves
66 their enzymatic activity and expands their substrate scope towards non-natural phenolic
67 starter units[20,21].

68 Recently also other flavonoids were directly produced with this simple pathway in microbial
69 cell factories by precursor feeding and pathway engineering. For instance, Dunstan *et al.*
70 established high titers for naringenin (484 mg/L from *p*-coumaric acid), eriodictyol (55mg/L
71 from caffeic acid), and homoeriodictyol (17mg/L from glycerol) with a semi-automated
72 metabolic engineering strategy in *Escherichia coli*[16]. Cui *et al*. obtained low final titers of
73 hesperetin (0.4mg/L) from isoferulic acid by constructing a recombinant *E. coli* strain that
74 expresses 4CL from *Oryza sativa* (Os4CL) and CHS from *Hordeum vulgare* (HvCHS)[22]. Since
75 efficient uptake of ferulic acid by *E. coli* has been demonstrated in other studies[23,24], it is most
76 likely that the low hesperetin titer in this study is related to overall poor enzymatic activity of
77 this enzyme combination, or the fact that ferulic acid and isoferulic acid are poor substrates
78 for these enzymes.

79 In this study, we set out to explore this enzyme combination further. We determined the
80 crystal structure of HvCHS in complex with CoA, CoA and naringenin and CoA, and eriodictyol.
81 Based on these structures, we designed mutant variants to increase the titers of the O-

82 methylated flavonoids homoeriodictyol, and hesperetin produced in *E. coli* from fed ferulic

83 acid and iso-ferulic acid.



84

85 **Figure 1.** Biosynthesis pathway of the four flavonoids discussed in the study. The natural

86 pathway product in plants is naringenin that is converted into the other flavonoids by tailoring

87 enzymes. In this study, eriodictyol, homoeriodictyol and hesperetin are directly produced by

88 feeding the respective precursors. (4CL 4-coumarate CoA ligase, CHS chalcone synthase, CHI

89 chalcone isomerase).

90 ## 2. Materials and methods

91 ### 2.1 Bacterial strains, primers, and plasmids

92 All the bacterial strains and plasmids used in this study are listed in table 1 and table 2. Primers
93 are listed in table S1. *E. coli* DH5a (New England Biolabs) was used for routine cloning and
94 pathway propagation, *E. coli* MG1655 K-12 (DE3) was used for protein expression and
95 fermentation, and *E. coli* BL21 (DE3) was used for protein purification[25]. The genes for PhCHS,
96 Chalcone synthase from *Petunia hybrida*; MsCHI, Chalcone isomerase from *Medicago sativa*;
97 Pc4CL, 4-coumarate--CoA ligase from *Petroselinum crispum*, HvCHS, Chalcone synthase from
98 *Hordeum vulgare*; Os4CL, 4-coumarate--CoA ligase from *Oryza sativa* were codon optimized
99 and synthesized from Integrate DNA Technology with restriction site introduced (Integrated
100 DNA Technologies, Coralville, Iowa, USA). The genes for PhCHS and HvCHS were cloned
101 between SacI and HindIII of pETDuet-1 and those for Pc4CL and Os4CL were cloned between
102 NdeI and XhoI of pETDuet-1 yielding pETDuet-HvCHS-Os4CL and pETDuet-PhCHS-Pc4CL. The
103 gene for MsCHI was cloned between NdeI and XhoI of pCDFDuet-1 and pETDuet-HvCHS
104 yielding pCDFDuet-MsCHI and pETDuet-HvCHS-MsCHI. The gene for Os4CL was cloned
105 between NdeI and XhoI of pCDFDuet-1 yielding pCDFDuet-Os4CL. The gene for HvCHS was
106 cloned between ScaI and HindIII of pET28a yielding pET28a-HvCHS. The gene for Os4CL was
107 cloned between NdeI and XhoI of pET28a yielding pET28a-Os4CL.

108 **Table 1.** List of plasmids and bacterial strains used in methylated flavonoids production strains.

Name Strains	Relevant Genotype	Reference
E. coli BL21(DE3)	F - ompT hsdSB(rB -mB -) gal dem (λIts857 indI Sam7 nin5 lavUV5-T7gene1)	Invitrogen
E. coli DH5a	supE44 Δ(lacZYA-argF)U196 (Φ80ΔlacZM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1	Invitrogen
E. coli K-12 MG1655 (DE3)	K-12 F^- λ^- *ilvG*$^-$ *rfb-50 rph-1*	Nielsen *et al*.[25]

Identifier	**Plasmid name**	**Backbone**	**Enzyme encoded**		**Source**
			MCS1	**MCS2**	
C1	pETDuet::PhCHS_Pc4CL	pETduet-1	PhCHS	Pc4CL	This study
C2	pCDFDuet::_MsCHI	pCDFduet-1	MsCHI	/	This study
C3	pETDuet::HvCHS_Os4CL	pETduet-1	HvCHS	Os4CL	This study
C4	pETDuet::PhCHS_Pc4CL(del V342)	pETduet-1	PhCHS	Pc4CL (del V342)	This study
C5	pETDuet::HvCHS (A228S, D231I, Q232P, L233G, D234V)_Os4CL	pETduet-1	HvCHS (A228S, D231I, Q232P, L233G, D234V)	Os4CL	This study

C6	pETDuet::PhCHS_Pc4CL(Q214A)	pETduet-1	PhCHS	Pc4CL (Q214A)	
C7	pETDuet::HvCHS_Os4CL (del V340)	pETduet-1	HvCHS	Os4CL (del V340)	This study
C8	pETDuet::HvCHS_Os4CL(Q212A)	pETduet-1	HvCHS	Os4CL (Q212A)	This study
C9	pETDuet::HvCHS_Os4CL(S242A)	pETduet-1	HvCHS	Os4CL (S242A)	This study
C10	pETDuet::HvCHS(A199T)_Os4CL	pETduet-1	HvCHS (A199T)	Os4CL	This study
C11	pETDuet::HvCHS(I267F)_Os4CL	pETduet-1	HvCHS (I267F)	Os4CL	This study
C12	pETDuet::PhCHS(T197A)_Pc4CL	pETduet-1	PhCHS (T197A)	Pc4CL	This study
C13	pETDuet::HvCHS(A228S)_Os4CL	pETduet-1	HvCHS (A228S)	Os4CL	This study
C14	pETDuet::HvCHS(D231I)_Os4CL	pETduet-1	HvCHS (D231I)	Os4CL	This study
C15	pETDuet::HvCHS(Q232P)_Os4CL	pETduet-1	HvCHS (Q232P)	Os4CL	This study
C16	pETDuet::HvCHS(L233G)_Os4CL	pETduet-1	HvCHS (L233G)	Os4CL	This study
C17	pETDuet::HvCHS(D234V)_Os4CL	pETduet-1	HvCHS (D234V)	Os4CL	This study
C18	pETDuet::HvCHS(Q232P, D234V)_Os4CL	pETduet-1	HvCHS (Q232P, D234V)	Os4CL	This study
C19	pETDuet::HvCHS(Q232P, D234V)_MsCHI	pETduet-1	HvCHS (Q232P, D234V)	MsCHI	This study
C20	pCDFDuet-1::Os4CL	pCDFDuet-1	Os4CL	/	This study
C21	pET28a::HvCHS	pET28a	HvCHS	/	This study
C22	pET28a::HvCHS(Q232P, D234V)	pET28a	HvCHS (Q232P, D234V)	/	This study
C23	pET28s::Os4CL	pET28a	Os4CL	/	This study

* MCS1 and 2, multiple cloning site (each with its own T7 promoter and terminator); PhCHS, CHS from *Petunia hybrida*; MsCHI, CHI from *Medicago sativa*; Pc4CL, 4CL from *Petroselinum crispum*; HvCHS, CHS from *Hordeum vulgare*; Os4CL, 4CL from *Oryza sativa*; codon optimized for expression in *E. coli.*

Table 2. List of *E. coli* MG1655 (DE3) strains used in fermentation experiments.

Identifier	Plasmid name	Enzymes expressed	Source
S1	C1, C2	PhCHS, Pc4CL, and MsCHI	This study
S2	C2, C3	HvCHS, Os4CL, and MsCHI	This study
S3	C2, C4	PhCHS, Pc4CL (del V342), and MsCHI	This study
S4	C2, C5	HvCHS (A228S, D231I, Q232P, L233G, D234V), Os4CL, and MsCHI	This study
S5	C2, C6	PhCHS, Pc4CL (Q214A), and MsCHI	This study
S6	C2, C7	HvCHS, Os4CL (del V340), and MsCHI	This study
S7	C2, C8	HvCHS, Os4CL (Q212A), and MsCHI	This study
S8	C2, C9	HvCHS, Os4CL (S242A), and MsCHI	This study
S9	C2, C10	HvCHS (A199T), Os4CL, and MsCHI	This study
S10	C2, C11	HvCHS (I267F), Os4CL, and MsCHI	This study
S11	C2, C12	PhCHS (T197A), Pc4CL, and MsCHI	This study
S12	C2, C13	HvCHS (A228S), Os4CL, and MsCHI	This study
S13	C2, C14	HvCHS (D231I), Os4CL, and MsCHI	This study
S14	C2, C15	HvCHS (Q232P), Os4CL, and MsCHI	This study
S15	C2, C16	HvCHS (L233G), Os4CL, and MsCHI	This study
S16	C2, C17	HvCHS (D234V), Os4CL, and MsCHI	This study
S17	C2, C19	HvCHS (Q232P, D234V), Os4CL, and MsCHI	This study
S18	C19, C20	HvCHS (Q232P, D234V), MsCHI, and Os4CL	This study

2.2 Synthesis of DNA template and protein mutagenesis

Mutagenesis libraries were generated using a modified QuikChange method[26]. The templates were digested with fast digest *Dpn*I overnight in 37°C incubators, then a PCR Clean-up kit was used to purify the QuikChange product. Subsequently, 1 µL of the purified DNA was transformed into *E. coli* DH5α. The plasmids were isolated with the QIAprep Spin Miniprep kit and were sequence verified.

119 ## 2.3 Transformation of *E. coli* for methylated flavonoid production

120 To generate the flavonoid producing *E. coli* strains s1-s18, the respective plasmid
121 combinations were co-transformed into *E. coli* MG1655 (DE3) electrocompetent cells. The
122 protocol for preparation of electrocompetent cell was from Green *et al*[27]. After transformation,
123 the cells were grown on selective LB agar (ampicillin and spectinomycin at 100 µg/mL) at 30°C
124 overnight. The next day, colonies were inoculated in 3 mL selective LB liquid medium in round-
125 bottom polystyrene tubes for cultivation at 30°C, 200 rpm overnight. The following day,
126 glycerol stocks were prepared from these cultures (50% (v/v) glycerol) and stored at -80°C
127 until further usage.

128 ## 2.4 Multiple sequence alignment

129 Protein sequences were aligned with Clustal Omega (1.2.4) with default settings[28]. The
130 alignment was visualized with AliView[29]. Active site residues of CHS were identified based on
131 Protein Data Bank (PDB) entry 1CGK and those of 4CL based on 5BSW.

132 ## 2.5 Fermentation conditions

133 For small-scale fermentation, the flavonoid pathway expressing strains were streaked in
134 duplicates from glycerol stocks into LB agar plates with ampicillin and spectinomycin
135 (100 µg/mL each). After incubation overnight at 30°C, single colonies were inoculated into a
136 starter culture of 3 mL of selective LB medium in round-bottom polystyrene tubes and grown
137 with shaking at 200 rpm at 30°C overnight. The next day, the cultures were diluted 1:100 into
138 3 mL seed cultures of modified MOPS medium[18] for 24 h at 30°C. These were used to inoculate
139 working cultures of 1 mL at an optical density at λ=600 nm (OD_{600}) of 0.05 in a 48 flower-
140 shaped well plate (m2p-lab, Germany) and incubated at 30°C, 900rpm in an Eppendorf
141 Thermomixer. IPTG (final concentration, 1 mM), precursors (final concentration, 1 mM), and
142 cerulenin (final concentration, 20 µg/mL) were added into the cell culture after 6 h and the
143 incubation was continued for 32 h.

144 For large-scale fermentation, two single colonies of the respective strains were inoculated into
145 starter cultures are described above. The next day, 1 mL of the starter culture was inoculated
146 into 50 mL of fresh LB medium. IPTG (final concentration 1 mM) was added to the culture
147 broths when the OD_{600} reached 0.4-0.6 for protein expression and the cultures were
148 successively incubated at 30 °C for an additional 3 h. The cells were collected by centrifugation
149 for 30 min at 3,700 rpm (Eppendorf 5920R, Germany) and 4 °C and then resuspended in 25 mL
150 of fresh modified MOPS medium which included 1 mM precursor, 1 mM IPTG, and 20 µg/mL
151 cerulenin for further 32 h fermentation[22].
152 Samples were taken for the quantification of OD_{600} and extracellular metabolites after 32h
153 fermentation[18]. Metabolites were prepared for HPLC and HPLC-MS analysis as described in
154 Dunstan *et al.*[16]. Samples were stored in -20 °C until analysis. OD_{600} was measured by an
155 absorbance plate reader (BMG Labtech, Germany) at 600 nm.

156 ## 2.6 Analysis and quantification of target compounds

157 The authentic standards for *p*-coumaric acid, caffeic acid, ferulic acid, isoferulic acid,
158 naringenin, eriodictyol, homoeriodictyol, hesperetin were purchased from Sigma Aldrich
159 (USA). The compounds used in *in vitro* experiments (coenzyme A hydrate, malonyl coenzyme
160 A tetralithium salt, and adenosine 5'-triphosphate (ATP) disodium salt hydrate) were
161 purchased from Sigma Aldrich (USA). Fatty acid synthase (FAS) inhibitor cerulenin was
162 purchased from Enzo life sciences (USA).
163 Samples from fermentation and *in vitro* turnover were analyzed by high performance liquid
164 chromatography (HPLC), with a Shimadzu LC-10AT system equipped with a SPD-20A
165 photodiode array detector (PDA). The samples were analyzed by 10µL injections and
166 separation over an Agilent Eclipse XDB-C18 (5 µm, 4.6×150 mm) column with a concentration
167 gradient (solution A: water + 0.1% trifluoroacetic acid (TFA), solution B: acetonitrile + 0.1%
168 TFA) at a flow rate of 1 mL/min. The following gradient was used: 15% solution B for 3 min,
169 15-90% solution B over 6 min; 90% solution B for 2 min; 90-15% solution B over 3 min, 15%
170 solution B for 4 min. *P*-coumaric acid, caffeic acid, ferulic acid, isoferulic acid, naringenin,

171 eriodictyol, homoeriodictyol, and hesperetin were identified by comparison to authentic
172 standards. The peak areas were integrated and converted to concentrations based on
173 calibration curves with the authentic standards (Fig. S1).
174 The identity of hesperetin was furthermore confirmed by HPLC coupled mass spectrometry
175 (HPLC-MS) with a Waters Acquity Arc UHPLC-MS equipped with a 2998 PDA, and a QDa single
176 quadrupole mass detector (Fig. S2). The samples were separated over an XBridge BEH C18 3.5
177 µm column with a concentration gradient (solution A: water + 0.1% formic acid, and solution
178 B: acetonitrile + 0.1% formic acid) at a flow rate of 0.25 mL/min (1 µL injections). The following
179 gradient was used: 5% solution B for 2 min, 5-90% solution B over 3 min; 90% solution B for 2
180 min; 5% solution B for 3 min.

181 ## 2.7 Protein expression, purification, and crystallization

182 **Expression**: *E. coli* BL21(DE3) harboring the plasmids pET28a::HvCHS, pET28a::HvCHS(Q232P,
183 D234V) or pET28a::Os4CL were inoculated in 3 mL LB (containing 100µg ml^{-1} kanamycin) for
184 overnight cultivation at 37°C, 200 rpm. The next day, 1 ml overnight cell culture was used to
185 inoculate 1 L fresh self-induction medium for further 20 h cultivation at 30°C, 200 rpm in 5 L
186 glass Erlenmeyer flask. The self-induction medium composition (1 L) consisted of 20 g tryptone,
187 5 g yeast, 5 g sodium chloride, 4.45 g disodium hydrogen phosphate dihydrate, 3.4 g
188 potassium dihydrogen phosphate, 6 g glycerol, 0.5 g glucose, 1.28 g lactose, and 100 µg ml^{-1}
189 Kanamycin.

190 **Purification**: All steps were performed with cooled buffers and at 4°C. Cells were harvested
191 by centrifugation at 3,700 rpm. The cell pellet was resuspended in 20 mL lysis buffer (50 mM
192 Tris-HCl, 300 mM NaCl, pH 7.6). The cells were disrupted by sonication for 4 × 40 s (with a
193 5 min rest interval between each cycle) at a 60 W output. The unbroken cells and debris were
194 removed by centrifugation (18,000 rpm for 1h). The supernatant was filtered through a
195 syringe filter (pore diameter 0.45 µm) and incubated with 3 mL Ni^{2+}-sepharose resin, which
196 had previously been equilibrated with lysis buffer, in a small column at 4°C for 18 h with
197 agitation. The unbound proteins were eluted from the column using gravity flow. The column

198 was first washed with lysis buffer (15 ml) and then with buffer A (30 ml, 50 mM Tris-HCl,
199 300 mM NaCl, 30 mM imidazole pH 7.6). Retained proteins were eluted with buffer B (5 ml,
200 50 mM Tris-HCl, 300 mM NaCl, 500 mM imidazole pH 7.6). Fractions were analyzed by
201 separation using sodium dodecyl sulfate polyacrylamide gel electrophoresis (4–12%
202 polyacrylamide) and staining with InstantBlue Coomassie Protein Stain (Abcam, UK). Fractions
203 containing chalcone synthase were combined and loaded onto a HiLoad 16/600 Superdex 200
204 pg column, which had previously been equilibrated with buffer C (180 ml, 10 mM HEPES, 50
205 mM NaCl buffer, 5% glycerol, and 2 mM DTT, pH 8.5). Elution was performed by running buffer
206 C across the column at 1 ml min^{-1} for 1.2 column volumes. Fractions were collected and
207 analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis. The purified enzyme
208 was concentrated with centrifugal devices with Omega membrane 30K (Pall, USA) and stored
209 at –80° C until further use.

210 **Crystallization:** Freshly prepared HvCHS was used in crystallization. The protein was
211 concentrated to 10 mg/ml in a buffer consisting of 10 mM HEPES pH7.5; 50 mM NaCl; 5% (v/v)
212 glycerol; 2 mM DTT. Screening for crystallization conditions was executed manually using
213 commercial sparse-matrix screening kits (JCSG plus; PACT premier; Morpheus and the PGA
214 screen; Molecular Dimensions Ltd.), by sitting drop vapor diffusion at 4°C. After 1-2 days, small
215 crystals were obtained in a drop containing 1 µl protein and 1 µl crystallization buffer (0.1 M
216 MES/Imidazole pH6.5; 0.03 M $MgCl_2$, 0.03 M $CaCl_2$; 20% (v/v) glycerol, 10% (v/v) PEG4000).
217 The crystallization buffer was optimized to a lower concentration of precipitants (16% (v/v)
218 glycerol and 8% (v/v) PEG4000) to get fewer and larger crystals (Fig. S3).

219 Crystals were harvested 3-4 days before diffraction using a nylon loop. Before flash-cooling in
220 liquid nitrogen, crystals were quickly dipped into a cryoprotectant, composed of the reservoir
221 buffer with an increased concentration of glycerol (32% (v/v)).

222 Some crystals were soaked in cryoprotectant containing either an additional 1 mM naringenin
223 or 5 mM eriodictyol for 24 h in order to obtain crystals of the enzyme-product complex.
224 Naringenin and eriodictyol were initially prepared as a 1 M stock in 100% DMSO.

225 **Data collection, Structure determination and refinement**

226 All diffraction data of HvCHS crystals were collected on beamlines P11 and P13 (operated by
227 EMBL Hamburg) of Petra III at DESY (Hamburg, Germany)[30,31].

228 Integration, space group determination, and scaling were carried out with XDSAPP and
229 Aimless in the CCP4 suite[32,33]. The structure of HvCHS complexed with CoA was determined by
230 molecular replacement with the monomer of chalcone synthase from *Arabidopsis thaliana*
231 (PDB: 6DXB) using the program Phaser[34]. Molecular replacement with the DIMPLE pipeline[35]
232 was performed to determine and initially refine the structures of HvCHS in complex with
233 naringenin and eriodictyol, utilizing the structure of HvCHS without ligands and waters as a
234 starting model.

235 All structures were iteratively refined using manual adjustment in Coot[36] and Refmac5[37].

236 **PDB deposition**

237 The structures of HvCHS in complex with CoA, CoA and naringenin, and CoA and eriodictyol
238 were deposited in the PDB under accession codes 8B32, 8B35, and 8B3C, respectively.

239 **Structure comparison and visualization:** The HvCHS structures were compared to known CHS
240 structures in the PDB and pairwise to each other with the Dali server[38]. Final structures were
241 visualized with PyMOL (Schrödinger, LLC).

242 ## 2.8 *In vitro* synthesis of feruloyl-CoA

243 Feruloyl-CoA was synthesized in 5 mL *in vitro* reaction with purified Os4CL enzyme following
244 established protocols[39,40]. The reaction mixture (purified enzyme (40 µg/ml), ferulic acid (400
245 µM), Coenzyme A (800 µM), ATP (2.5 mM), and $MgCl_2$ (5 mM) in potassium phosphate buffer
246 (50 mM, pH 7.4)) was incubated at 30°C in the dark, with mixing at 200 rpm overnight.
247 Reaction product was analyzed by HPLC and then aliquoted and stored at −20°C for further
248 experiments.

249 ## 2.9 *In vitro* enzymatic assay

250 The *in vitro* enzymatic assays with HvCHS were performed in 50 µL reaction mixtures
251 composed of feruloyl-CoA (2.5, 5, 10, 25, 50, 100 µM), malonyl-CoA (300 µM), and HvCHS
252 variant (50 nM) in 100 mM phosphate buffer (pH 7.4). The individual reactions were started
253 by adding the enzyme, incubated at 37°C without mixing, and quenched at different time
254 points (2.5, 5, and 7.5 min) by adding an equal volume of methanol with 0.1% formic acid to
255 quench the reaction. The samples were centrifuged at 15,000 rpm for 10 min, and the
256 supernatant was used for HPLC analysis. The integrated peak areas of the product were
257 converted into concentrations based on a calibration curve with the authentic standard, and
258 plotted against time (Fig. S4). Apparent initial velocities were determined by linear regression
259 over the early time points before 10% substrate conversion was reached. The apparent initial
260 velocities of the biological triplicates were then plotted against the substrate concentrations
261 and the resulting curves were fitted with the Michaelis-Menten equation in GraphPad prism.
262 The full statistical analysis is provided in Table S2.

263 # 3. Results

264 Based on the results of previous studies on flavonoid production in *E. coli*, we chose two
265 combinations of CHS and 4CL as a starting point for this study: PhCHS from *Petunia hybrida*
266 and Pc4CL from *Petroselinum crispum*, previously shown to yield high titers of naringenin[15],
267 and HvCHS from *Hordeum vulgare* and Os4CL from *Oryza sativa*, previously shown to accept
268 O-methylated cinnamic acids to form homoeriodictyol and hesperetin[22]. We cloned the genes
269 into co-expression plasmids and transformed them into *E. coli* MG1655 (DE3) (Table 2). In an
270 initial fermentation experiment at 1 mL scale, with fed ferulic acid, and isoferulic acid (1 mM
271 final concentration), we observed that the combination of HvCHS and Os4CL (s2) yielded up
272 to two times higher titers of homoeriodictyol than the other enzyme combination (s1) and
273 that hesperetin was only produced by s2 (Fig. S5). Intrigued by this result, we wondered if
274 mutagenesis of the 4CL enzymes to better accommodate the O-methylated substrates could
275 further increase the titers of hesperetin and homoeriodictyol. Based on a previous

276 mutagenesis study showing that the deletion of V341 in the *Nicotiana tabacum* 4CL allows for
277 binding of double methylated sinapinic acid[41], we deleted the corresponding residues in Pc4CL
278 (V342) and Os4CL (V340). Furthermore, based on the crystal structure of the *Nicotiana*
279 *tabacum* 4CL in complex with feruloyl-CoA (PDB: 5BSW)[41], we hypothesized that mutating
280 Q213 and S243 into alanine might provide more space in the substrate binding pocket to
281 accommodate the O-methyl group. In Pc4CL, residue 243 is already an alanine. We screened
282 the new 4CL mutants in the recombinant flavonoid pathway against the two methylated
283 substrates and did not observe remarkable differences (Fig. S6) although the final
284 homoeriodictyol titer with s6 (Os4CL delV340) is around 1.5 times higher than s2 for
285 production (Fig. S6).

286 Thus, we proceeded with investigating the role of CHS in substrate selection by determining
287 the crystal structures of HvCHS with the ligands CoA, naringenin, and eriodictyol by X-ray
288 crystallography and by site-directed mutagenesis**.**

289 3.1 Crystal structure of chalcone synthase from *Hordeum vulgare*

290 The initial crystallization experiments with purified HvCHS yielded high-quality crystals that
291 diffracted to 1.7 Å (Table 3). We solved the phase problem with molecular replacement using
292 PDB: 6DXB [42] without the ligand. The refined structure (PDB: 8B32) shows high structural
293 similarity with several CHS structures in the PDB based on a Dali search[43] (Table S3). The
294 highest Z-scores were determined for CHS1 from *Oryza sativa*, CHS from *Freesia hybrida*,
295 (mutant variants of) CHS from *Medicago sativa* (MsCHS) and CHS1 from *Glycine max* (L.).
296 These enzymes all share an amino acid sequence identity of more than 70% and adopt the
297 typical CHS fold with the αβαβα pseudo-symmetric thiolase motif in the C-terminal domain
298 and the α-helical motif in the N-terminal domain[44] (Fig. 2A). In the refined structure, we
299 observed additional electron density in the substrate binding cleft separating the N- and the
300 C-terminal domain, which we interpreted as CoA. The putative CoA molecule occupies the
301 area of the enzyme previously identified as the malonyl-CoA binding pocket[44] and our

302 **Table 3.** Diffraction data collection, structure determination, and refinement statistics. Values
303 in brackets refer to the highest resolution shell.

Structure name		HvCHS+CoA	HvCHS+CoA, naringenin	HvCHS+CoA, eriodictyol
PDB entry		**8B32**	**8B35**	**8B3C**
Space group		P21 21 21	P21 21 21	P21 21 21
Unit cell parameter	a,b,c [Å]	a = 77.81	a = 73.91	a = 77.88
		b = 98.96	b = 95.98	b = 97.82
		c = 138.11	c = 138.34	c = 137.93
	α, β, γ [°]	α = 90.00	α = 90.00	α = 90.00
		β = 90.00	β = 90.00	β = 90.00
		γ = 90.00	γ = 90.00	γ = 90.00
Resolution, Å		45.79-1.70 (1.73 - 1.70)	47.99 - 2.00 (2.05- 2.00)	48.91 - 2.00 (2.04 - 2.00)
Number of observations		1556039 (8829)	886223 (8504)	921816 (8357)
Completness,%		99.8 (97.7)	99.9 (99.3)	100 (98.7)
Multiplicity		13.2 (13.7)	13.2 (13.5)	12.8 (12.5)
Rmerge		0.07 (1.366)	0.09 (0.476)	0.08 (1.016)
Average I/σ,I		18.3 (2.6)	19.8 (6.4)	17.0 (3.1)
CC(1/2)		0.999 (0.885)	0.999 (0.979)	0.999 (0.947)
Refinement statistics				
No. reflections, all/free		117356/5895	67037/3335	71816/3554
Rfactor		0.177	0.159	0.198
Rfree		0.205	0.193	0.234
No. protein atoms		5859	5804	5838
No. ligand atoms		96	136	138
No. water atoms		401	367	164
Average B-factors, $Å^2$				
Protein		34.58	30.26	50.26
Ligands		55.21	56.41	83.47
Water		40.78	36.32	48.50
rmsd from ideal values				
Bond lengths, Å		0.0129	0.0119	0.0094
Bond angles, °		1.757	1.658	1.503
Ramachandran plot				
Favored, %		96.97	96.31	97.11
Allowed, %		2.50	3.03	2.63

304 interpretation of the electron density agrees well with the model of the ligand in the CoA-
305 bound structure of MsCHS (PDB: 1BQ6) (Fig. 2B).

306 To gain further insight into the binding pocket of the cinnamic acid starter unit, we soaked the
307 crystals of HvCHS with naringenin, eriodictyol, homoeriodictyol, and hesperetin. While the
308 soaking with the O-methylated flavonoids impaired crystal diffraction, we were able to
309 determine the structures of the naringenin- (PDB: 8B35) and eriodictyol-bound enzyme (PDB:
310 8B3C) at 2Å resolution. In both structures, the electron density for CoA and the products are

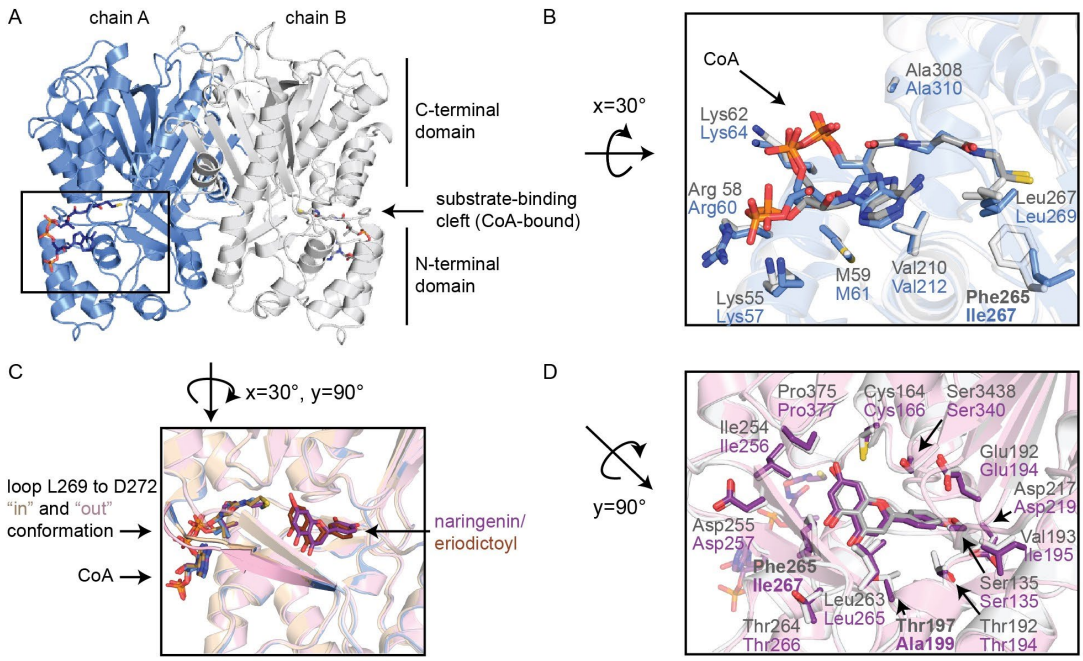


Figure 2. Crystal structures of HvCHS. A) Overview of HvCHS fold with the two monomers shown in grey and blue. B) CoA binding pocket of HvCHS in complex with CoA (blue) compared to the structure of MsCHS in complex with CoA (PDB: 1BQ6, grey). C) Overlay of all HvCHS structures (CoA-bound (blue), naringenin-bound (pink), eriodictoyl-bound (yellow). D) product binding pocket in HvCHS with naringenin bound (pink) compared to the structure of MsCHS in complex with naringenin (PDB: 1CGK, grey). Atom coloring in ligand representation: oxygen=red, nitrogen=blue, phosphorous=orange, carbon=color matches the cartoon color of the corresponding structure.

well-defined and the position of the bound products agrees well with the MsCHS structure (PDB: 1CGK). All three HvCHS structures are highly congruent, even in residues lining the active site. The only difference can be seen in the β-sheet and loop from L269 to D272 (Fig 2C). Based on the electron density map, this loop appears to adopt two distinct conformations in all three crystal structures. However, while in the CoA-bound and the CoA/eriodictoyl-bound structures, the "in"-conformation has much higher occupancy and was therefore chosen for the final model, the "out"-conformation is dominant in the naringenin-bound structure. This agrees well with the conformation of this region in the naringenin-bound structure of MsCHS (1CGK, Fig 2D). This may be related to the slight difference in the angle of the A- and C-ring of naringenin and eriodictyol in the active site. The B-ring and all residues surrounding it, show identical conformations in both product-bound structures.

331 When comparing the structures and amino acid sequences of MsCHS and HvCHS, it becomes
332 apparent that the residues in the substrate binding channel, and especially the cinnamoyl-
333 binding pocket are highly conserved (Fig 2D, Fig. S7). The only differences are residue 197 in
334 MsCHS, where a threonine is located compared to an alanine in HvCHS (residue 199), and
335 residue 265, where a phenylalanine is located compared to an isoleucine in HvCHS (residue
336 267). Residue 197 is close to the C-ring of naringenin and may influence substrate binding.
337 Mutation of residue 265 into valine was previously shown to decrease the efficiency of MsCHS
338 two-fold but it did not alter the substrate scope of the enzyme[21]. Further away from the active
339 site, we also noticed a small stretch of amino acids (228-234) with a markedly different
340 sequence in HvCHS than in MsCHS and PhCHS. These residues are located on the surface of
341 the enzyme in a loop that follows a β-strand, the opposite end of which forms part of the
342 active site. Despite the dissimilarity in sequence, this loop adopts the same conformation in
343 the MsCHS and HvCHS crystal structures (Fig. 28).

344 Based on our analysis of the sequence and structure of HvCHS, we decided to generate several
345 mutant variants of HvCHS and PhCHS to probe the effect of these minor sequence differences
346 on the enzyme performance for the production of methylated flavonoids *in vivo*.

347 3.2 Screening of CHS mutant variants for the production of methylated flavonoids

348 We constructed PhCHS T197A, HvCHS A199T, HvCHS I267F, and a HvCHS loop variant, where
349 the sequence of the surface loop (228-234) is replaced by the sequence of the PhCHS loop.
350 We transformed the resulting plasmids into *E. coli* to generate the flavonoid producing strains
351 s4, s9, and s10 for the HvCHS variants and s11 for the PhCHS T197A variant. We performed
352 small-scale fermentations (1 mL) of these strains feeding 1 mM ferulic acid or isoferulic acid
353 as precursors, and sampled 32 h after induction of enzyme expression (Fig. 3A). The final
354 optical densities (OD_{600}) of all cultures were similar (ranging from 0.23 to 0.25) suggesting that
355 the mutant variants do not affect the growth or viability of *E. coli*. Both point mutations in
356 HvCHS (s9, s10) did not alter the final titers of homoeriodictyol compared to s2, yet the A199T

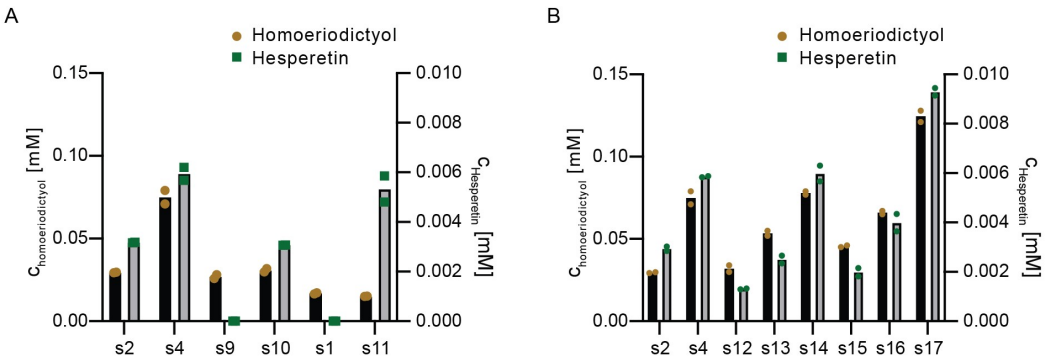


Figure 3. Small-scale fermentation of *E. coli* flavonoid producing strains expressing CHS variants. Titers of hesperetin (green) and homoeriodictyol (brown) were determined 32 h post induction of enzyme expression and addition of the precursors (1 mM ferulic acid or isoferulic acid). The experiment was performed in duplicate. A) Comparison of *E. coli* flavonoid producing strains expressing CHS variants: s1 PhCHS wildtype, s11 PhCHS T197A, s2 HvCHS wildtype, s4 loop variant, s9 HvCHS A199T, and s10 HvCHS I267F. B) Comparison of *E. coli* flavonoid producing strains expressing HvCHS variants: s2 wildtype, s4 loop variant, s12-s16 single point mutation variants, s17 double point mutation variant.

mutation completely eliminated hesperetin formation in s9. The analogous point mutation in PhCHS (T197A, s11) for the first time enabled hesperetin production with this enzyme, while the homoeriodictyol titer from this strain is comparable to the strain expressing the wildtype protein (s1). Most interestingly, strain s4 expressing the HvCHS loop variant shows increased titers of both methylated flavonoids with final titers reaching 0.07 and 0.006 mM. These titers are about two-fold higher than the final titer achieved with HvCHS wildtype, and the hesperetin titer is comparable to the one achieved with the PhCHS T197A variant (s11). This suggests that the presence of alanine in this position in the active site enables the binding of isoferulic acid but it does not have a strong effect on ferulic acid binding. Furthermore, the residues in the surface loop appear to also influence the *in vivo* performance of the HvCHS enzyme, yet with an unknown mechanism.

In order to further probe the importance of the individual residues of the surface loop for the enhanced *in vivo* performance of the HvCHS loop variant (A228S, D231I, Q232P, L233G, D234V), we next constructed single mutants for those five sites via site-directed mutagenesis. We transformed the resulting plasmids into *E. coli* (s12-s16) and compared the final titers of

381 fermentations with these strains to the ones expressing HvCHS wildtype and the loop variant
382 (s2 and s4, respectively). Compared to s4, only s14 and s16 yielded similar or higher product
383 titers (Fig. 3B). S12 produced lower and s13 and s15 similar titers as s2 expressing wildtype
384 HvCHS. Therefore, we combined the two best point mutations into a double mutant HvCHS
385 (Q232P, D234V) and transformed the resulting plasmid into *E. coli* (s17). This strain yields
386 higher titers of the methylated flavonoids than all other strains, with final homoeriodictyol
387 titers 2-fold higher, and final hesperetin titers 3-fold higher than with strain s2 expressing
388 wildtype HvCHS (Fig. 3B).

389 ### 3.3 Further optimization of flavonoid production

390 To test if we could further increase the titers of methylated flavonoids with our HvCHS variant,
391 we explored a different plasmid configuration. Santos *et al*. reported that it was optimal for
392 naringenin production to clone CHS into multiple cloning site (MCS) 1 of pETDuet-1, MsCHI
393 into MCS2 and 4CL into MCS1 of pCDFduet-1[15]. Thus, we constructed expression plasmids c19
394 and c20 with our best mutant variant according to this design and co-transformed them into
395 *E. coli* (s18). We performed a fermentation at larger scale in shake flasks to see how the titers
396 for methylated flavonoids from s18 compare to those of s2 and s17 (Fig. 4A). We found that
397 the final titers of both products from s18 are comparable to those obtained with s2 (0.24 mM
398 for homoeriodictyol and 0.011 mM for hesperetin). The final titer of homoeriodictyol and
399 hesperetin are highest with s17, 0.33 mM and 0.016 mM, respectively. We next examined the
400 production of two unmethylated flavonoids, naringenin and eriodictyol, from their respective
401 cinnamic acid precursors with our best strain in shake flasks. After 32 h, we achieved titers of
402 41 mg/L and 45 mg/L, respectively. Overall, our best strain s17 produces similar or higher titers
403 of the four flavonoids compared to the study of Cui *et al.*[22].

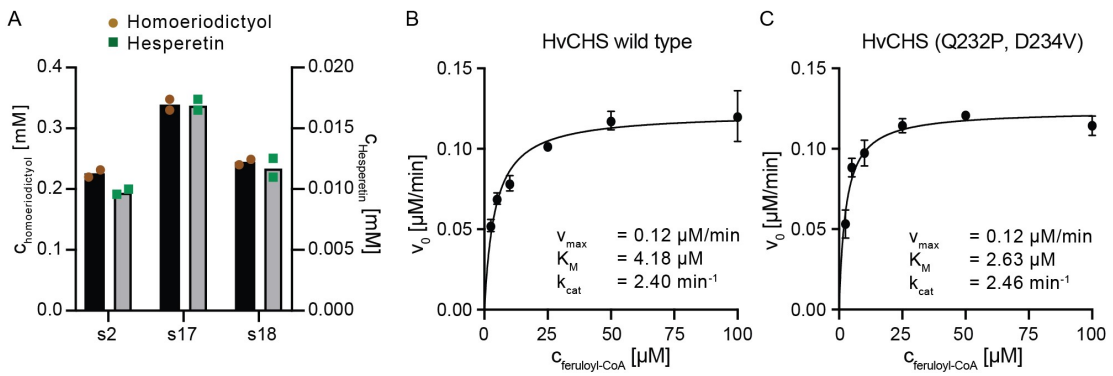


Figure 4. A) Comparison of HvCHS wildtype and double mutant for methylated flavonoid production in a large-scale fermentation. 1 mM ferulic acid and isoferulic acid were added as precursors. Samples were taken after 32 h fermentation and analyzed by HPLC. Each experiment was duplicated. HvCHS variants *in vitro* kinetic data. B-C) Kinetic characterization of HvCHS wildtype and HvCHS double mutant for homoeriodictyol formation. Apparent initial velocities were determined at 5, 10, 25, 50, and 100 µM substrate concentration (feruloyl-CoA). Data points represent mean +/- SD, n=3. The full statistical report for non-linear regression is shown in Table S2.

3.4 Enzymatic assay for HvCHS variants

Intrigued by the observation that the surface loop residues of HvCHS have an impact on the enzyme's *in vivo* performance, we endeavored to further examine this phenomenon *in vitro*. We expressed and purified HvCHS wildtype and double mutant (Q232P, D234V) with protein yields of 48 and 93 mg protein/ L culture, respectively. This suggests that the double mutant has a higher expression level, is more soluble, or more stable during purification than the wildtype enzyme. Given that homoeriodictyol is the main product *in vivo*, we performed steady-state kinetic assays for wildtype and double mutant using feruloyl-CoA as a substrate at a fixed malonyl-CoA concentration. Plotting the apparent initial reaction velocities over the substrate concentration allowed us to fit the data with the Michaelis-Menten equation (Fig. 4B/C, Fig. S4, Table S2). The apparent kinetic parameters v_{max} and k_{cat} are virtually the same for both HvCHS variants, whereas there is a 1.6 fold difference in the K_M. This indicates that the HvCHS (Q232P, D234V) variant exhibits a higher affinity towards feruloyl-CoA than the wildtype and is in good agreement with our results from homoerioditiyol production *in vivo*. In summary, the enhanced biosynthesis of flavonoids (naringenin, eriodictyol,

428 homoeriodictyol, and hesperetin) in *E. coli* expressing the HvCHS double mutant (Q232P,
429 D234V) can likely be attributed to the combined effects of a higher protein expression level,
430 and a higher affinity to methylated substrates.

431 4. Discussion and Conclusions

432 Hesperetin and homoeriodictyol are valuable natural O-methylated flavonoids which display
433 bioactivity for treating human disease, such as neuroinflammation, memory impairment and
434 carcinoid cancer[45,46]. Recently, Cui *et al*. obtained hesperetin and homoeriodictyol at a low
435 titer in engineered *E. coli* from fed isoferulic acid and ferulic acid[22]. The enzymes used in that
436 study had previously been shown to have high substrate promiscuity *in vitro*[47,48] and were
437 then used *in vivo* for the first time[22]. To become industrially relevant, the yield and
438 productivity of such a pathway must be dramatically increased and enzymes must be highly
439 selective for the O-methylated substrates to allow for the use of low-cost substrates. Thus, to
440 facilitate further enzyme engineering efforts, we determined the crystal structures of one of
441 the enzymes of this pathway, HvCHS, in complex with its products and analyzed a multiple
442 sequence alignment with the most similar structures in the PDB. Since all CHS enzymes are
443 highly conserved overall, especially in the substrate binding pockets, we did not see any
444 obvious reasons why HvCHS would accept O-methylated precursors and other CHS enzymes
445 would not. However, we identified three areas of interest for rational enzyme engineering –
446 two amino acids in the substrate binding pocket that are markedly different from the
447 consensus sequence (A199 and I267 in HvCHS) and a surface exposed loop that is connected
448 to the active site through a β-strand and has more polar residues than in the other CHS
449 sequences. Indeed, mutating the conserved Thr at position 197 in PhCHS enabled this enzyme
450 to produce hesperetin for the first time. However, replacing I267 with the canonical Phe did
451 not alter the performance of HvCHS. Surprisingly, mutating two of the polar or charged
452 residues of the surface loop into hydrophobic amino acids increased the final titers of
453 methylated flavonoids *in vivo*. We were able to show that this is likely attributable to a boost
454 in protein abundance and an increase in affinity for the methylated cinnamic acid.

455 The final titers that we achieved for both methylated flavonoids in a larger scale experiment,
456 exceed the previously reported titers by 2- and 10-fold for homoeriodictyol and hesperetin
457 production. Compared to homoeriodictyol, hesperetin production is still very low, mainly due
458 to the substrate preference of HvCHS. In the future, this substrate preference can be further
459 shifted by directed evolution or rational engineering based on our crystal structure. In
460 particular, in light of our surprising findings about the surface loop distant to the active site, it
461 is worth exploring other positions in the multiple sequence alignment that are markedly
462 different in HvCHS compared to other CHS enzymes. Alternatively, as more than 1000 putative
463 plant CHSs have now been predicted via computational tools, it is also interesting to
464 investigate the substrate scope and catalytic activity of these new enzymes to possibly replace
465 the workhorse CHS enzymes with ones that are better suited for large-scale production of
466 hesperetin and other flavonoids. Furthermore, an alternative strategy could involve
467 employing host strains that possess an ample supply of malonyl-CoA for enhanced flavonoid
468 synthesis.

Abbreviations and Nomenclature

4CL, 4-coumarate: CoA ligase

CHI, chalcone isomerase

CHS, chalcone synthase

CoA, Co-enzyme A

HPLC, High Performance Liquid Chromatography

MS, mass spectrometry

TFA, trifluoroacetic acid

Supporting Information

Gene and primer sequences, results of Dali analysis, statistical analysis of steady-state kinetics, calibration curves for product quantification by HPLC, HPLC-MS identification of fermentation products, micrograph of protein crystals, time progress curves for steady-state kinetics, product titers of fermentations with strains s1, s2, s3, s5, s6, and s7 expressing 4CL variants, sequence comparison of CHS enzymes, illustration of HvCHS surface loop (PDF).

Acknowledgement

The authors acknowledge DESY (Hamburg, Germany) and EMBL Hamburg for the provision of the PETRA III beamline (P11 and P13). KH, BP, SH and WJQ thank Serj Koshian for designing the initial set of mutant enzymes during his summer internship, Dr. Robbert Hans Cool for providing training in protein purification, and Nika Sokolova for help with troubleshooting the steady-state kinetics.

Funding Sources

BP, LZ and SH are supported by promotion scholarships from the Chinese Scholarship Council (202008420246, 202006320070, and 201806300121). KH is grateful for funding from the European Union's Horizon 2020 research and innovation program under the Marie

Skłodowska-Curie grant agreement No 893122. WQ was supported by SNN, Netherlands 338 (grant number OPSNN0315).

495 ## References

496 1. Panche, A. N., Diwan, A. D. & Chandra, S. R. Flavonoids: an overview. *J Nutr Sci* **5**, e47
497 (2016).

498 2. Kočevar, N., Glavač, I. & Kreft, S. FLAVONOIDS. in *Chemotaxonomy of Flowering*
499 *Plants* vol. 58 523–619 (MQUP, 2007).

500 3. Samanta Amalesh, Das Gouranga, D. S. K. Roles of flavonoids in Plants. *Int J Pharm Sci*
501 *Tech* **6**, 12–35 (2011).

502 4. Wang, Y., Chen, S. & Yu, O. Metabolic engineering of flavonoids in plants and
503 microorganisms. *Appl Microbiol Biotechnol* **91**, 949–956 (2011).

504 5. Sajid, M., Channakesavula, C. N., Stone, S. R. & Kaur, P. Synthetic biology towards
505 improved flavonoid pharmacokinetics. *Biomolecules* **11**, (2021).

506 6. Małgorzata Brodowska, K. Natural flavonoids: classification, potential role, and
507 application of flavonoid analogues. *Eur J Biol Res* **7**, 108–123 (2017).

508 7. Koirala, N., Thuan, N. H., Ghimire, G. P., Thang, D. Van & Sohng, J. K. Methylation of
509 flavonoids: Chemical structures, bioactivities, progress and perspectives for
510 biotechnological production. *Enzyme Microb Technol* **86**, 103–116 (2016).

511 8. Saquib, Q. *et al.* Anticancer efficacies of persicogenin and homoeriodictyol isolated
512 from *Rhus retinorrhoea*. *Process Biochemistry* **95**, 186–196 (2020).

513 9. Shen, T. *et al.* Homoeriodictyol protects human endothelial cells against oxidative
514 insults through activation of Nrf2 and inhibition of mitochondrial dysfunction. *Vascul*
515 *Pharmacol* **109**, 72–82 (2018).

516 10. Kheradmand, E., Hajizadeh Moghaddam, A. & Zare, M. Neuroprotective effect of
517 hesperetin and nano-hesperetin on recognition memory impairment and the
518 elevated oxygen stress in rat model of Alzheimer's disease. *Biomedicine &*
519 *Pharmacotherapy* **97**, 1096–1101 (2018).

520 11. Hermawan, A., Putri, H. & Utomo, R. Y. Comprehensive bioinformatics study reveals
521 targets and molecular mechanism of hesperetin in overcoming breast cancer
522 chemoresistance. *Mol Divers* **24**, 933–947 (2020).

523 12. Dong, W. *et al.* A dual character of flavonoids in influenza A virus replication and
524 spread through modulating cell-autonomous immunity by MAPK signaling pathways.
525 *Sci Rep* **4**, 7237 (2014).

526 13. Lin, C.-W. *et al.* Anti-SARS coronavirus 3C-like protease effects of *Isatis indigotica* root
527 and plant-derived phenolic compounds. *Antiviral Res* **68**, 36–42 (2005).

528 14. Cai, C.-Y. *et al.* Analogues of xanthones——Chalcones and bis-chalcones as α-
529 glucosidase inhibitors and anti-diabetes candidates. *Eur J Med Chem* **130**, 51–59
530 (2017).

531 15. Santos, C. N. S., Koffas, M. & Stephanopoulos, G. Optimization of a heterologous
532 pathway for the production of flavonoids from glucose. *Metab Eng* **13**, 392–400
533 (2011).

534 16. Dunstan, M. S. *et al.* Engineering *Escherichia coli* towards de novo production of
535 gatekeeper (2S)-flavanones: naringenin, pinocembrin, eriodictyol and
536 homoeriodictyol. *Synth Biol* **5**, 1–11 (2020).

537 17. Koopman, F. *et al.* De novo production of the flavonoid naringenin in engineered
538 *Saccharomyces cerevisiae*. *Microb Cell Fact* **11**, 155 (2012).

539 18. Dinh, C. V. & Prather, K. L. J. Development of an autonomous and bifunctional
540 quorum-sensing circuit for metabolic flux control in engineered *Escherichia coli*.
541 *Proceedings of the National Academy of Sciences* **116**, 25562–25568 (2019).

542 19. Sun, W. *et al.* Molecular and Biochemical Analysis of Chalcone Synthase from *Freesia*
543 *hybrida* in Flavonoid Biosynthetic Pathway. *PLoS One* **10**, e0119054 (2015).

544 20. Bhan, N., Cress, B. F., Linhardt, R. J. & Koffas, M. Expanding the chemical space of
545 polyketides through structure-guided mutagenesis of *Vitis vinifera* stilbene synthase.
546 *Biochimie* **115**, 136–143 (2015).

547 21. Jez, J. M. *et al.* Expanding the biosynthetic repertoire of plant type III polyketide
548 synthases by altering starter molecule specificity. *Proc Natl Acad Sci U S A* **99**, 5319–
549 24 (2002).

550 22. Cui, H., Song, M. C., Lee, J. Y. & Yoon, Y. J. Microbial production of O-methylated
551 flavanones from methylated phenylpropanoic acids in engineered *Escherichia coli*. *J*
552 *Ind Microbiol Biotechnol* **46**, 1707–1713 (2019).

553 23. Li, B., He, X., Zhang, S., Chang, S. & He, B. Efficient synthesis of 4-O-β-d-
554 glucopyranosylferulic acid from ferulic acid by whole cells harboring
555 glycosyltransferase GTBP1. *Biochem Eng J* **130**, 99–103 (2018).

556 24. Yoon, S.-H. *et al.* Production of vanillin from ferulic acid using recombinant strains of
557 *Escherichia coli*. *Biotechnology and Bioprocess Engineering* **10**, 378–384 (2005).

558 25. Nielsen, D. R., Yoon, S.-H., Yuan, C. J. & Prather, K. L. J. Metabolic engineering of
559 acetoin and meso-2, 3-butanediol biosynthesis in *E. coli*. *Biotechnol J* **5**, 274–84
560 (2010).

561 26. Kunkel, T. A. Rapid and efficient site-specific mutagenesis without phenotypic
562 selection. *Proceedings of the National Academy of Sciences* **82**, 488–492 (1985).

563 27. Green, M. R. & Sambrook, J. Transformation of *Escherichia coli* by Electroporation.
564 *Cold Spring Harb Protoc* **2020**, pdb.prot101220 (2020).

565 28. Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022.
566 *Nucleic Acids Res* **50**, W276–W279 (2022).

567 29. Larsson, A. AliView: a fast and lightweight alignment viewer and editor for large
568 datasets. *Bioinformatics* **30**, 3276–3278 (2014).

569 30. Meents, A. *et al.* Development of an in-vacuum x-ray microscope with cryogenic
570 sample cooling for beamline P11 at PETRA III. in *Proceedings of SPIE - The*
571 *International Society for Optical Engineering TA - TT* - (ed. Lai, B.) vol. 8851 88510K
572 (2013).

573 31. Cianci, M. *et al.* P13, the EMBL macromolecular crystallography beamline at the low-
574 emittance PETRA III ring for high- and low-energy phasing with variable beam
575 focusing. *J Synchrotron Radiat* **24**, 323–332 (2017).

576 32. Krug, M., Weiss, M. S., Heinemann, U. & Mueller, U. XDSAPP: A graphical user
577 interface for the convenient processing of diffraction data using XDS. *J Appl*
578 *Crystallogr* **45**, 568–572 (2012).

579 33. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta*
580 *Crystallogr D Biol Crystallogr* **67**, 235–242 (2011).

581 34. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674
582 (2007).

583 35. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta*
584 *Crystallogr D Biol Crystallogr* **67**, 235–242 (2011).

585 36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of
586 Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).

587 37. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal
588 structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–67 (2011).

589 38. Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res* **50**,
590 W210–W215 (2022).

591 39. Kong, D., Li, S. & Smolke, C. D. Discovery of a previously unknown biosynthetic
592 capacity of naringenin chalcone synthase by heterologous expression of a tomato
593 gene cluster in yeast. *Sci Adv* **6**, eabd1143 (2020).

594 40. Rautengarten, C., Baidoo, E., Keasling, J. D. & Scheller, H. V. A simple method for
595 enzymatic synthesis of unlabeled and radiolabeled hydroxycinnamate-CoA. *Bioenergy*
596 *Res* **3**, 115–122 (2010).

597 41. Li, Z. & Nair, S. K. K. Structural Basis for Specificity and Flexibility in a Plant 4-
598 Coumarate:CoA Ligase. *Structure* **23**, 2032–2042 (2015).

599 42. Liou, G., Chiang, Y.-C., Wang, Y. & Weng, J.-K. Mechanistic basis for the evolution of
600 chalcone synthase catalytic cysteine reactivity in land plants. *J Biol Chem* **293**, 18601–
601 18612 (2018).

602 43. Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res* **50**,
603 W210–W215 (2022).

604 44. Ferrer, J.-L., Jez, J. M., Bowman, M. E., Dixon, R. A. & Noel, J. P. Structure of chalcone
605 synthase and the molecular basis of plant polyketide biosynthesis. *Nat Struct Biol* **6**,
606 775–784 (1999).

607 45. Muhammad, T., Ikram, M., Ullah, R., Rehman, S. & Kim, M. Hesperetin, a Citrus
608 Flavonoid, Attenuates LPS-Induced Neuroinflammation, Apoptosis and Memory
609 Impairments by Modulating TLR4/NF-κB Signaling. *Nutrients* **11**, 648 (2019).

610 46. Zarebczan, B., Pinchot, S. N., Kunnimalaiyaan, M. & Chen, H. Hesperetin, a potential
611 therapy for carcinoid cancer. *The American Journal of Surgery* **201**, 329–333 (2011).

612 47. Lee, Y.-J. *et al.* Enzymatic Synthesis of Phenolic CoAs Using 4-Coumarate:coenzyme A
613 Ligase (4CL) from Rice. *Bull Korean Chem Soc* **28**, 365–366 (2007).

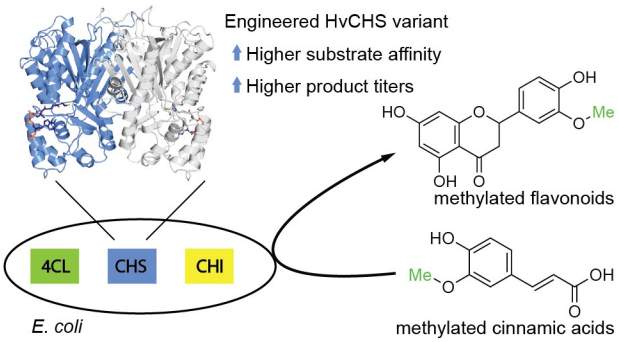
614 48. Christensen, A. B., Gregersen, P. L., Schröder, J. & Collinge, D. B. A chalcone synthase
615 with an unusual substrate preference is expressed in barley leaves in response to UV
616 light and pathogen attack. *Plant Mol Biol* **37**, 849–857 (1998).

617

Author contributions

KH and BP conceived the study with contributions from LZ, SH, RO, WJQ and MG; BP and SH cloned plasmids and performed fermentations; BP and KH analyzed biochemical data; LZ and RO performed crystallization and diffraction experiments; LZ, RO and MG determined and refined the crystal structures; KH and BP wrote the manuscript with contributions from LZ, SH, RO, WJQ and MG; all authors have read and approved the final version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

For table of contents only