Heng *et al.*
Biotechnology for Biofuels and Bioproducts (2024) 17:131
https://doi.org/10.1186/s13068-024-02578-2

Biotechnology for Biofuels and Bioproducts

RESEARCH

Open Access

Expanding the biosynthesis spectrum of hydroxy fatty acids: unleashing the potential of novel bacterial fatty acid hydratases

Yu Chyuan Heng[1*], Garrett Wei Jie Wong[1,2] and Sandra Kittelmann[1*]

Abstract

Background Hydroxy fatty acids represent an emerging class of compounds with promising applications in the chemical, medicinal and functional food sectors. The challenges associated with their chemical synthesis have spurred exploration of biological synthesis as an alternative route, particularly through the use of fatty acid hydratases. Fatty acid hydratases catalyse the regioselective addition of a hydrogen atom and a hydroxyl group from a water molecule to the carbon–carbon *cis*-double bond of unsaturated fatty acids to form hydroxy fatty acids. Despite having been discovered in the early 1960s, previous research has primarily focused on characterizing single fatty acid hydratase variants with a limited range of substrates. Comprehensive studies that systematically examine and compare the characteristics of multiple variants of fatty acid hydratases are still lacking.

Results In this study, we employed an integrated bioinformatics workflow to identify 23 fatty acid hydratases and characterized their activities against nine unsaturated fatty acid substrates using whole-cell biotransformation assays. Additionally, we tested a dual-protein system involving two fatty acid hydratases of distinct regioselectivity and demonstrated its suitability in enhancing the biosynthesis of di-hydroxy fatty acids.

Conclusions Our study demonstrates that fatty acid hydratases can be classified into three subtypes based on their regioselectivity and provides insights into their preferred substrate structures. These understandings pave ways for the design of optimal fatty acid hydratase variants and bioprocesses for the cost-efficient biosynthesis of hydroxy fatty acids.

Keywords Fatty acid hydratases, Hydroxy fatty acids, Unsaturated fatty acids, Regioselectivity, Substrate selectivity, Dual-protein system

Introduction

Hydroxy fatty acids (HFAs) are saturated and unsaturated fatty acids with one or more hydroxyl functional groups attached to carbon atoms along the carbon chain. They are essential chemical compounds for producing biodegradable polymers, lubricants, emulsifiers, drugs, cosmetic ingredients, and flavours. Some HFAs and their derivatives have also been shown to possess beneficial biological activities [1–8], indicating suitability for application in the medical and functional foods sectors. Major advantages of biological over chemical synthesis of HFAs are the higher regio- and stereoselectivity of the reaction.

*Correspondence:
Yu Chyuan Heng
yuchyuan.heng@sg.wilmar-intl.com
Sandra Kittelmann
sandra.kittelmann@sg.wilmar-intl.com
[1] Wilmar International Limited, 28 Biopolis Road, Singapore 138568, Singapore
[2] Present Address: Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, 14 Medical Drive, Singapore 117599, Singapore

Among the different classes of fatty acid-hydroxylation enzymes, fatty acid hydratases (FAHs; EC 4.2.1.53) have drawn considerable attention due to their wide substrate range as well as their widespread distribution in food-safe-microorganisms such as lactobacilli. FAHs are flavin adenine dinucleotide (FAD)-containing enzymes that catalyse the regioselective addition of a hydrogen atom and a hydroxyl group from a water molecule to the carbon–carbon *cis*-double bond of unsaturated fatty acids to form HFAs. Those FAHs that hydroxylate the *cis*-9 double bond of oleic acid to produce 10-hydroxystearic acid (10-HSA) are traditionally known as oleate hydratases and have been studied since the early 1960s [9, 10]. It has since been demonstrated that a vast variety of FAHs with different positional selectivity exist in nature. These enzymes are spread across numerous bacterial phyla and are classified into 11 homologous families (HFams) based on protein sequence similarity [11]. FAHs within the same HFam, however, may exhibit a completely different positional selectivity, and thus, substrate and product scopes. For example, despite within-HFam sequence identity of > 62%, HFam 2 contains characterized FAHs that hydroxylate the *cis*-double bond/s of C14 to C22 unsaturated fatty acids at position 9, at position 12, or at both positions, to produce mono- or di-HFAs (Table S1). Additionally, FA-HY1 from *Lactobacillus acidophilus*, has also been shown to hydroxylate unsaturated fatty acids with a double bond at *cis*-11, -13, or -14 positions [12]. Notably, the regioselectivity of FAHs can be improved through protein engineering approaches, as demonstrated by Zhang et al., who shifted the regioselectivity of FA-HY1 towards the *cis*-14 position, generating a single 15-HFA product from eicosapentaenoic acid [13]. Despite the numerous studies conducted on FAHs, most have primarily focused on characterizing single variants with a limited range of substrates. Comprehensives studies that examine and compare the characteristics of multiple FAHs are still lacking.

In this study, we employed an integrated bioinformatics workflow and whole-cell biotransformation assays to identify and characterize a collection of FAHs exhibiting the ability to hydroxylate a range of monounsaturated (MUFA) and polyunsaturated fatty acid (PUFA) substrates to form specific HFAs. Furthermore, we proposed a dual-protein system and demonstrated its suitability in enhancing the synthesis of di-HFAs from PUFAs. Our study demonstrates that FAHs can be classified into three different subtypes based on regioselectivity and HFA product scope. Our findings on the specific substrate preferences and characteristics of different FAH subtypes and protein variants contribute to deepening the understanding of the sequence-structure–function relationship within this enzyme class.

Results and discussion

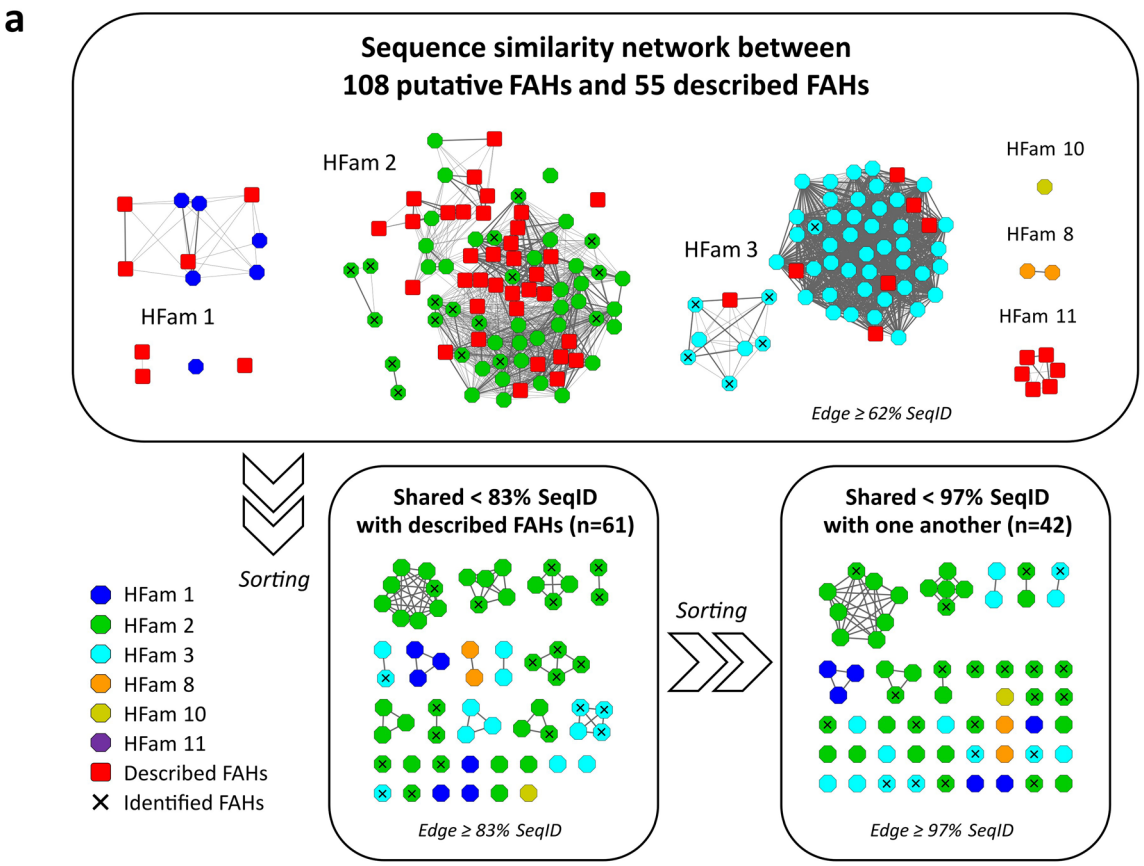
FAH identification through in silico gene mining and sequence similarity analysis

FAHs belong to a family of highly conserved proteins known as myosin-cross-reactive antigen (MCRA; PF06100). To identify FAH homologs, we conducted a homology search using the consensus protein sequence of the hidden Markov model (HMM) of the MCRA protein family as a query against a database of draft genomes curated from 139 in-house bacterial strains [14–16]. This search yielded 108 putative FAHs, which were subsequently reduced to 61 proteins after excluding those that shared > 83% sequence identity with 55 previously described FAHs (Fig. 1a). We further refined the list to 42 proteins by retaining only a single entry from each group of closely related putative FAHs that shared > 97% sequence identity to one another. From these, we selected 21 putative FAHs for screening of fatty acid-hydroxylation activity, alongside two putative FAHs that shared > 90% sequence identity to previously described FAHs (Tables 1 and S2).

Among the 23 selected putative FAHs, 17 were categorised under HFam 2, and the remaining 6 under HFam 3. In the phylogenetic tree, putative FAHs belonging to the same HFam generally aggregated into the same clades, with proteins from related taxa clustering closer to one another than to those from distantly related taxa (Fig. 1b). The lengths of the 23 putative FAHs ranged from 557 to 645 amino acids. They were identified in 22 distinct bacterial species, including an isolate of a novel species designated as *Limosilactobacillus* sp. WILCCON 0055. Notably, two putative FAHs (WIL0055-HY35.1

(See figure on next page.)

Fig. 1 In silico gene mining and sequence similarity analysis. **a** Sequence similarity analysis compared 108 putative FAHs with 55 described FAHs (and among themselves) to identify 42 proteins that shared < 83% sequence identity with the described FAHs and < 97% sequence identity to one another. In total, 21 of these putative FAHs, alongside two putative FAHs that shared > 90% sequence identity to the described FAHs, were selected for the screening of fatty acid-hydroxylation activity. **b** Phylogenetic tree based on protein sequences showing the relationship between the 23 FAHs. The maximum-likelihood tree was inferred under the LG + G4 protein substitution model (optimal log-likelihood of -11608.2102) using IQ-TREE, rooted by midpoint-rooting, and visualized using iTOL. Ultrafast bootstrap values based on 1,000 replications are indicated at branch points. The scale bar denotes 0.1 substitutions per amino acid position. SeqID, sequence identity



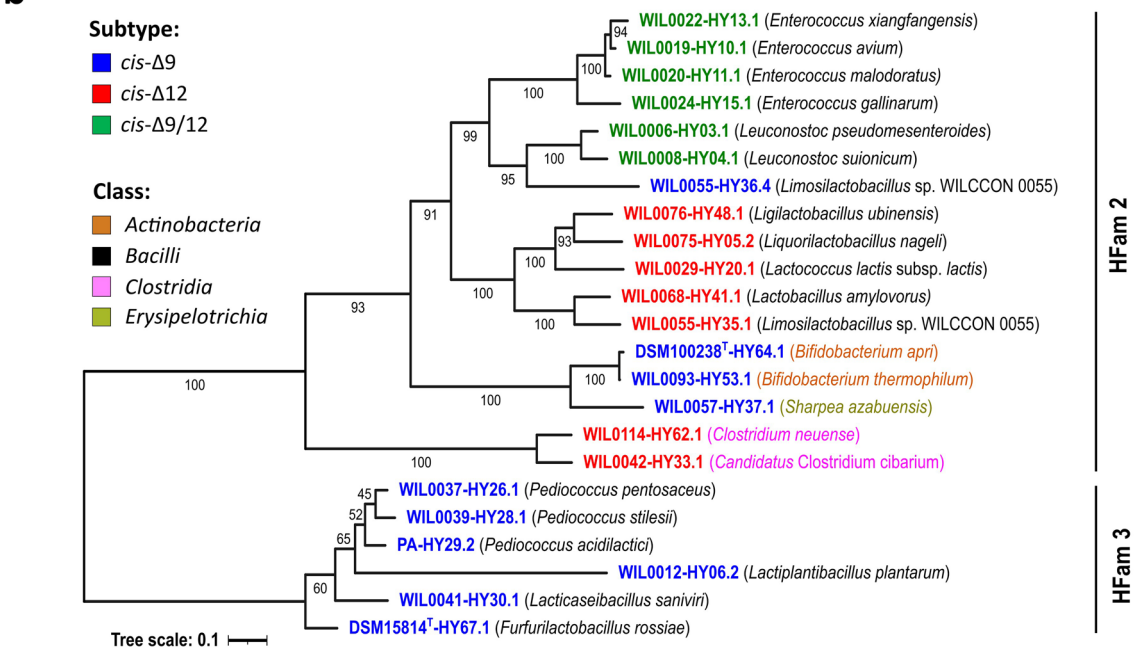


Fig. 1 (See legend on previous page.)

Table 1 List of 23 FAHs identified and characterized in this study sorted according to the families of bacterial strains from which they were derived

Phylum; class; order; family	Strain	FAH	Subtype	HFam	Amino acid length
Actinomycetota; *Actinobacteria*; *Bifidobacteriales*; *Bifidobacteriaceae*	*Bifidobacterium apri* DSM 100238^{T}	DSM100238^{T}-HY64.1	*cis*-Δ9	2	622
	Bifidobacterium thermophilum WILCCON 0093	WIL0093-HY53.1	*cis*-Δ9	2	642
Bacillota; *Bacilli*; *Lactobacillales*; *Enterococcaceae*	*Enterococcus avium* WILCCON 0019	WIL0019-HY10.1	*cis*-Δ9/12	2	590
	Enterococcus gallinarum WILCCON 0024	WIL0024-HY15.1	*cis*-Δ9/12	2	590
	Enterococcus malodoratus WILCCON 0020	WIL0020-HY11.1	*cis*-Δ9/12	2	590
	Enterococcus xiangfangensis WILCCON 0022	WIL0022-HY13.1	*cis*-Δ9/12	2	590
Bacillota; *Bacilli*; *Lactobacillales*; *Lactobacillaceae*	*Furfurilactobacillus rossiae* DSM 15814^{T}	DSM15814^{T}-HY67.1	*cis*-Δ9	3	559
	Lacticaseibacillus saniviri WILCCON 0041	WIL0041-HY30.1	*cis*-Δ9	3	558
	Lactiplantibacillus plantarum WILCCON 0012	WIL0012-HY06.2	*cis*-Δ9	3	564
	Lactobacillus amylovorus WILCCON 0068	WIL0068-HY41.1	*cis*-Δ12	2	590
	Leuconostoc pseudomesenteroides WILCCON 0006	WIL0006-HY03.1	*cis*-Δ9/12	2	590
	Leuconostoc suionicum WILCCON 0008	WIL0008-HY04.1	*cis*-Δ9/12	2	590
	Ligilactobacillus ubinensis WILCCON 0076	WIL0076-HY48.1	*cis*-Δ12	2	589
Bacillota; *Bacilli*; *Lactobacillales*; *Lactobacillaceae*	*Limosilactobacillus* sp. WILCCON 0055	WIL0055-HY35.1	*cis*-Δ12	2	591
		WIL0055-HY36.4	*cis*-Δ9	2	605
	Liquorilactobacillus nagelii WILCCON 0075	WIL0075-HY05.2	*cis*-Δ12	2	589
	Pediococcus pentosaceus WILCCON 0037	WIL0037-HY26.1	*cis*-Δ9	3	557
	Pediococcus stilesii WILCCON 0039	WIL0039-HY28.1	*cis*-Δ9	3	571
	Pediococcus acidilactici PA	PA-HY29.2	*cis*-Δ9	3	559
Bacillota; *Bacilli*; *Lactobacillales*; *Streptococcaceae*	*Lactococcus lactis* subsp. *lactis* WILCCON 0029	WIL0029-HY20.1	*cis*-Δ12	2	589
Bacillota; *Clostridia*; *Eubacteriales*; *Clostridiaceae*	*Clostridium neuense* WILCCON 0114	WIL0114-HY62.1	*cis*-Δ12	2	564
	Candidatus Clostridium cibarium WILCCON 0042	WIL0042-HY33.1	*cis*-Δ12	2	564
Bacillota; *Erysipelotrichia*; *Erysipelotrichales*; *Coprobacillaceae*	*Sharpea azabuensis* WILCCON 0057	WIL0057-HY37.1	*cis*-Δ9	2	645

FAH subtype, HFam classification, and amino acid length are provided

and WIL0055-HY36.4) were found in this novel isolate. Importantly, none of the putative FAHs from these bacterial strains have been previously identified, cloned, or expressed in recombinant *E. coli* strains to test for their fatty acid-hydroxylation potential.

Screening of FAHs for fatty acid-hydroxylation activity

Upon selection, the nucleotide sequences encoding the 23 putative FAHs were cloned into the pET-15b vector, under the regulation of the *lac* operon (Fig. S1). The resulting expression plasmids were transformed into *E. coli* strain Acella, generating recombinant strains Acella/pET-15b-XXX (XXX refers to the abbreviation of each FAH). Recombinant FAH expression was induced by culturing the transformant strains in an autoinduction medium containing 1.5% (w/v) lactose and verified through sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE).

To screen for fatty acid-hydroxylation activity, a whole-cell biotransformation assay was performed. The induced recombinant *E. coli* strains were incubated with 100 g L^{-1} Wilfarin OA-7075, a mixture comprising 81.0% oleic acid (OA; C18:1 $\Delta^{9(Z)}$) and 10.1% linoleic acid (LA: C18:2 $\Delta^{9(Z),12(Z)}$), at 40 °C for 48 h. The resulting reaction mixtures were analysed by liquid chromatography-tandem mass spectrometry (LC–MS/MS) to determine their fatty acid composition. Based on this screening assay, a putative FAH was identified as active if it produced HFA(s) from one or both the provided M/PUFA substrates (Fig. 2a). Our results indicated that all 23 putative FAHs produced HFA(s) under the tested expression and screening conditions (Fig. 2b), thereby confirming their activity as FAHs.

The identified FAHs could be broadly classified into three different subtypes based on their regioselectivity and HFA product scope. We designated FAHs that produced 10-hydroxystearic acid (10-HSA) from OA and (12*Z*)-10-hydroxyoctadec-12-enoic acid (10-HOE) from LA as the *cis*-Δ9 subtype, pertaining to hydroxylation of the *cis*-9 double bond. Conversely, FAHs that produced

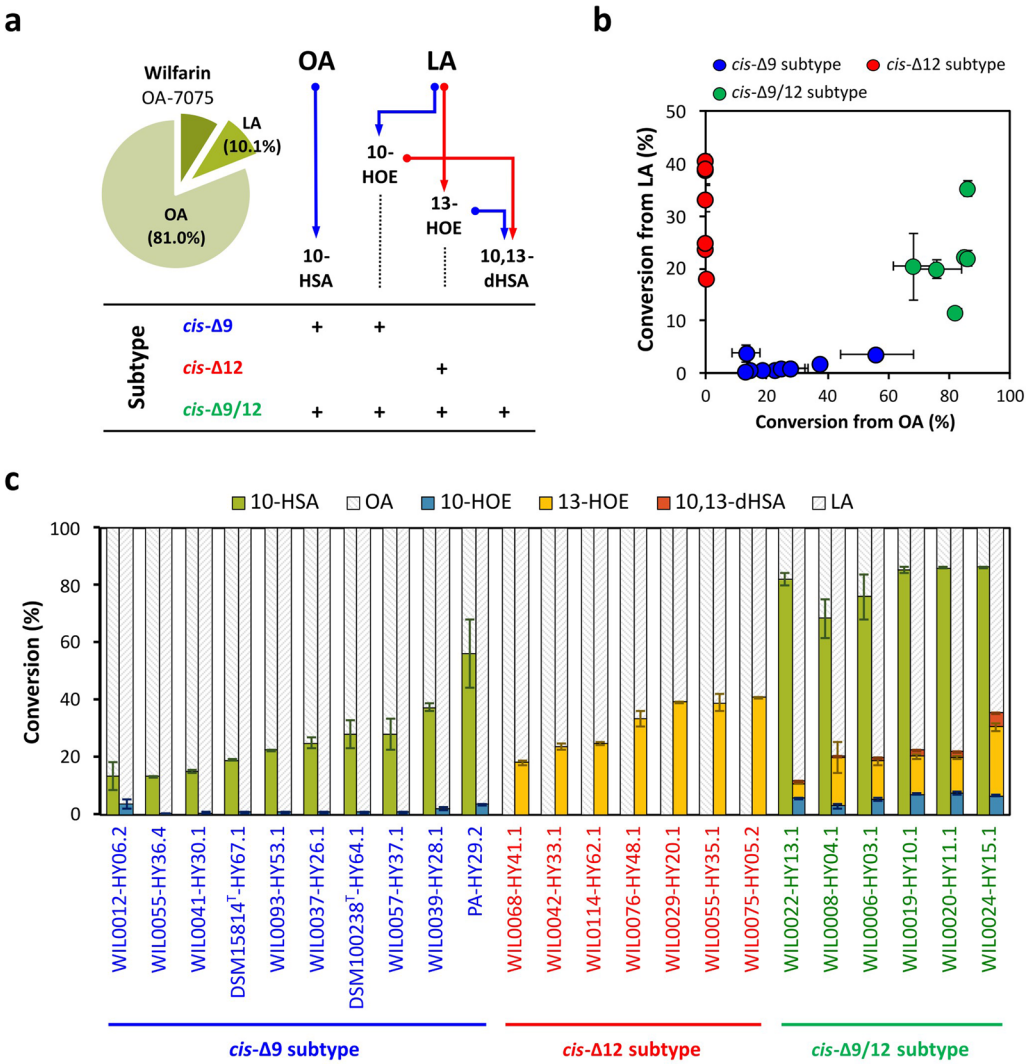


Fig. 2 Screening of FAHs for fatty acid-hydroxylation activity. **a** The identified FAHs were classified into three different subtypes based on the profile of HFA produced (regioselectivity) from OA and LA in Wilfarin OA-7075. **b** Total conversion from OA and LA by FAHs of different subtypes. The graph shows the mean percentage conversion ± standard deviations of two independent replicates after 48 h of incubation at 40 °C. **c** Percentage conversion of OA into 10-HSA, and LA into 10-HOE, 13-HOE and 10,13-dHSA by FAHs of different subtypes. Bars show the mean percentage conversion ± standard deviations of two independent replicates after 48 h of incubation at 40 °C. For fatty acid abbreviations, refer to Table 2

(9*Z*)-13-hydroxyoctadec-9-enoic acid (13-HOE) from LA were categorised as the *cis*-Δ12 subtype, attributing to hydroxylation of the *cis*-12 double bond. FAHs that hydroxylated both the *cis*-9 and *cis*-12 double bonds, producing 10-HSA from OA, and 10-HOE, 13-HOE and 10,13-dihydroxystearic acid (10,13-dHSA) from LA, were classified as the *cis*-Δ9/12 subtype (Fig. 3c and f).

Among the 23 identified FAHs, ten belonged to the *cis*-Δ9 subtype, seven to the *cis*-Δ12 subtype, and six to the *cis*-Δ9/12 subtype (Fig. 2c). FAHs of the *cis*-Δ9 subtype were found across HFam 2 and 3, while those of the *cis*-Δ12 and *cis*-Δ9/12 subtypes were exclusive to HFam 2 (Fig. 1b). Notably, the two FAHs identified in the novel *Limosilactobacillus* sp. WILCCON 0055 isolate varied considerably in their activity and positional selectivity. WIL0055-HY35.1 was selective to the *cis*-12 double bond while WIL0055-HY36.4 was selective to the *cis*-9 double bond. In general, we observed relatively higher levels of production of 10-HSA and 10-HOE for FAHs of the *cis*-Δ9/12 subtype than the *cis*-Δ9 subtype, whereas the level of production of 13-HOE was higher for FAHs of the *cis*-Δ12 subtype than the *cis*-Δ9/12 subtype (Fig. 2c).

Regioselectivity, substrate selectivity, and product scopes

FAHs have previously been reported to hydroxylate not only the *cis*-9 and *cis*-12 double bonds of M/PUFAs with

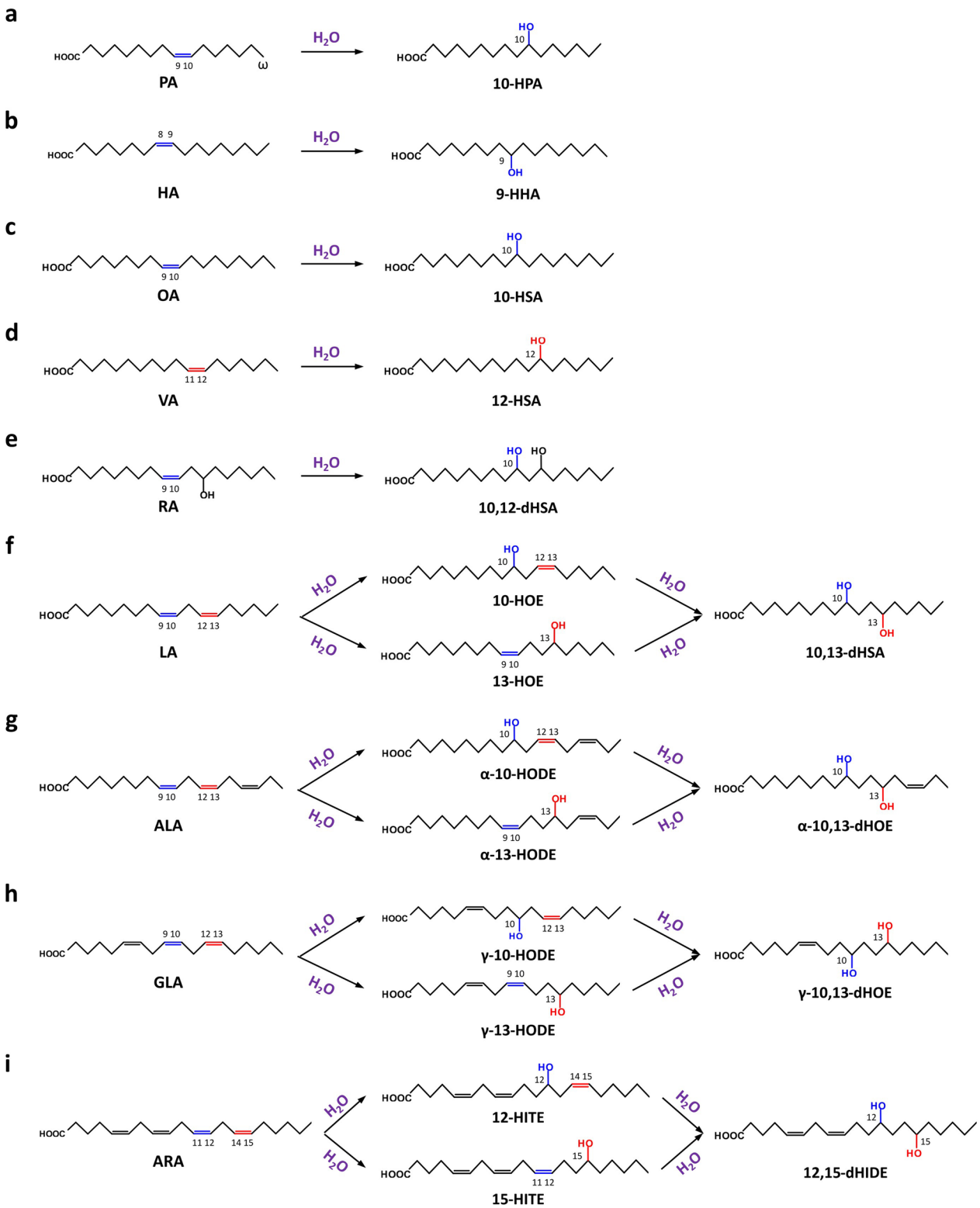


Fig. 3 Hydroxylation of unsaturated fatty acids by FAHs of different subtypes. **a**–**e** Hydroxylation of MUFAs, **f**–**i** hydroxylation of PUFAs. *Cis*-double bonds targeted by FAHs of the *cis*-Δ9 subtype are indicated in blue, whereas *cis*-double bonds targeted by FAHs of the *cis*-Δ12 subtype are indicated in red. FAHs of the *cis*-Δ9/12 subtype could hydroxylate all *cis*-double bonds targeted by the other two subtypes. For fatty acid abbreviations, refer to Table 2

hydrocarbon chains other than C18. To better understand their regioselectivity, substrate selectivity, and product scopes, we selected nineteen FAHs representing the top performers of each subtype and subjected their induced recombinant *E. coli* cells to another series of whole-cell biotransformation reactions involving a panel of nine M/PUFA substrates (30 mM final concentration). In addition to OA and LA, we tested palmitoleic acid (PA; C16:1 $\Delta^{9(Z)}$), (8*Z*)-heptadec-8-enoic acid (HA; C17:1 $\Delta^{8(Z)}$), vaccenic acid (VA; C18:1 $\Delta^{11(Z)}$), ricinoleic acid (RA; C18:1 $\Delta^{9(Z)}$, 12-OH), α-linolenic acid (ALA; C18:3 $\Delta^{9(Z),12(Z),15(Z)}$), γ-linolenic acid (GLA; C18:3 $\Delta^{6(Z),9(Z),12(Z)}$) and arachidonic acid (ARA; C20:4 $\Delta^{5(Z),8(Z),11(Z),14(Z)}$). This expanded substrate panel allowed examination of hydroxylation at eight different *cis*-double bond positions in various even-chain (C16, C18, and C20) M/PUFAs, as well as an odd-chain (C17) MUFA (Table 2 and Fig. 3).

The results revealed consistent profiles of HFAs produced from the nine different M/PUFA substrates across FAHs within each subtype (Fig. 4). All six FAHs of the *cis*-Δ9 subtype were regioselective towards the *cis*-9 double bond, converting PA, OA, RA, LA, ALA, and GLA into the corresponding 10-HFA products. Although there was slight variation in substrate selectivity between FAHs, a general trend of higher conversion efficiency was observed for OA (84.4–91.7%), followed by RA (35.0–91.8%), PA (29.6–99.6%), GLA (9.8–55.0%), ALA (1.6–46.7%) and LA (2.1–28.7%). This suggested a preference of FAHs of the *cis*-Δ9 subtype for MUFAs over PUFAs, and C18 over C16 hydrocarbon chain length. Additionally, we observed the conversion of HA to 9-hydroxyheptadecanoic acid (9-HHA; 1.6–46.9%) and ARA to (5*Z*,8*Z*,14*Z*)-12-hydroxyicosa-5,8,14-trienoic acid (12-HITE; 0.1–0.9%), suggesting hydroxylation of the *cis*-8 and *cis*-11 double bond of a C17 MUFA and a C20 PUFA, respectively.

Similarly, all seven FAHs of the *cis*-Δ12 subtype exhibited regioselectivity towards the *cis*-12 double bond, producing 13-HFA products from LA, ALA, and GLA. Despite subtle variation between FAHs, a general trend of higher conversion efficiency was noted in the order of LA (62.4–71.5%), GLA (13.1–77.7%) and ALA (2.2–35.0%). This indicated a preference of FAHs of the *cis*-Δ12 subtype for PUFAs of the C18 hydrocarbon chain

Table 2 List of fatty acid abbreviations, their common and IUPAC names, and chemical formulas. NA, not available

Abbreviation	Common name	IUPAC name	Chemical formula
PA	Palmitoleic acid	(9*Z*)-hexadec-9-enoic acid	C16:1 $\Delta^{9(Z)}$
10-HPA	10-hydroxypalmitic acid	10-hydroxyhexadecanoic acid	C16:0, 10-OH
HA	NA	(8*Z*)-heptadec-8-enoic acid	C17:1 $\Delta^{8(Z)}$
9-HHA	NA	9-hydroxyheptadecanoic acid	C17:0, 9-OH
OA	Oleic acid	(9*Z*)-octadec-9-enoic acid	C18:1 $\Delta^{9(Z)}$
10-HSA	10-hydroxystearic acid	10-hydroxyoctadecanoic acid	C18:0, 10-OH
VA	Vaccenic acid	(11*Z*)-octadec-11-enoic acid	C18:1 $\Delta^{11(Z)}$
12-HSA	12-hydroxystearic acid	12-hydroxyoctadecanoic acid	C18:0, 12-OH
RA	Ricinoleic acid	(9*Z*)-12-hydroxyoctadec-9-enoic acid	C18:1 $\Delta^{9(Z)}$, 12-OH
10,12-dHSA	10,12-dihydroxystearic acid	10,12-dihydroxyoctadecanoic acid	C18:0, 10,12-OH
LA	Linoleic acid	(9*Z*,12*Z*)-octadeca-9,12-dienoic acid	C18:2 $\Delta^{9(Z),12(Z)}$
10-HOE	NA	(12*Z*)-10-hydroxyoctadec-12-enoic acid	C18:1 $\Delta^{12(Z)}$, 10-OH
13-HOE	NA	(9*Z*)-13-hydroxyoctadec-9-enoic acid	C18:1 $\Delta^{9(Z)}$, 13-OH
10,13-dHSA	10,13-dihydroxystearic acid	10,13-dihydroxyoctadecanoic acid	C18:0, 10,13-OH
ALA	α-linolenic acid	(9*Z*,12*Z*,15*Z*)-octadeca-9,12,15-trienoic acid	C18:3 $\Delta^{9(Z),12(Z),15(Z)}$
α-10-HODE	NA	(12*Z*,15*Z*)-10-hydroxyoctadeca-12,15-dienoic acid	C18:2 $\Delta^{12(Z),15(Z)}$, 10-OH
α-13-HODE	NA	(9*Z*,15*Z*)-13-hydroxyoctadeca-9,15-dienoic acid	C18:2 $\Delta^{9(Z),15(Z)}$, 13-OH
α-10,13-dHOE	NA	(15*Z*)-10,13-dihydroxyoctadec-15-enoic acid	C18:1 $\Delta^{15(Z)}$, 10,13-OH
GLA	γ-linolenic acid	(6*Z*,9*Z*,12*Z*)-octadeca-6,9,12-trienoic acid	C18:3 $\Delta^{6(Z),9(Z),12(Z)}$
γ-10-HODE	NA	(6*Z*,12*Z*)-10-hydroxyoctadeca-6,12-dienoic acid	C18:2 $\Delta^{6(Z),12(Z)}$, 10-OH
γ-13-HODE	NA	(6*Z*,9*Z*)-13-hydroxyoctadeca-6,9-dienoic acid	C18:2 $\Delta^{6(Z),9(Z)}$, 13-Oh
γ-10,13-dHOE	NA	(6*Z*)-10,13-dihydroxyoctadec-6-enoic acid	C18:1 $\Delta^{6(Z)}$, 10,13-OH
ARA	Arachidonic acid	(5*Z*,8*Z*,11*Z*,14*Z*)-icosa-5,8,11,14-tetraenoic acid	C20:4 $\Delta^{5(Z),8(Z),11(Z),14(Z)}$
12-HITE	NA	(5*Z*,8*Z*,14*Z*)-12-hydroxyicosa-5,8,14-trienoic acid	C20:3 $\Delta^{5(Z),8(Z),14(Z)}$, 12-OH
15-HITE	NA	(5*Z*,8*Z*,11*Z*)-15-hydroxyicosa-5,8,11-trienoic acid	C20:3 $\Delta^{5(Z),8(Z),11(Z)}$, 15-OH
12,15-dHIDE	NA	(5*Z*,8*Z*)-12,15-dihydroxyicosa-5,8-dienoic acid	C20:2 $\Delta^{5(Z),8(Z)}$, 12,15-OH

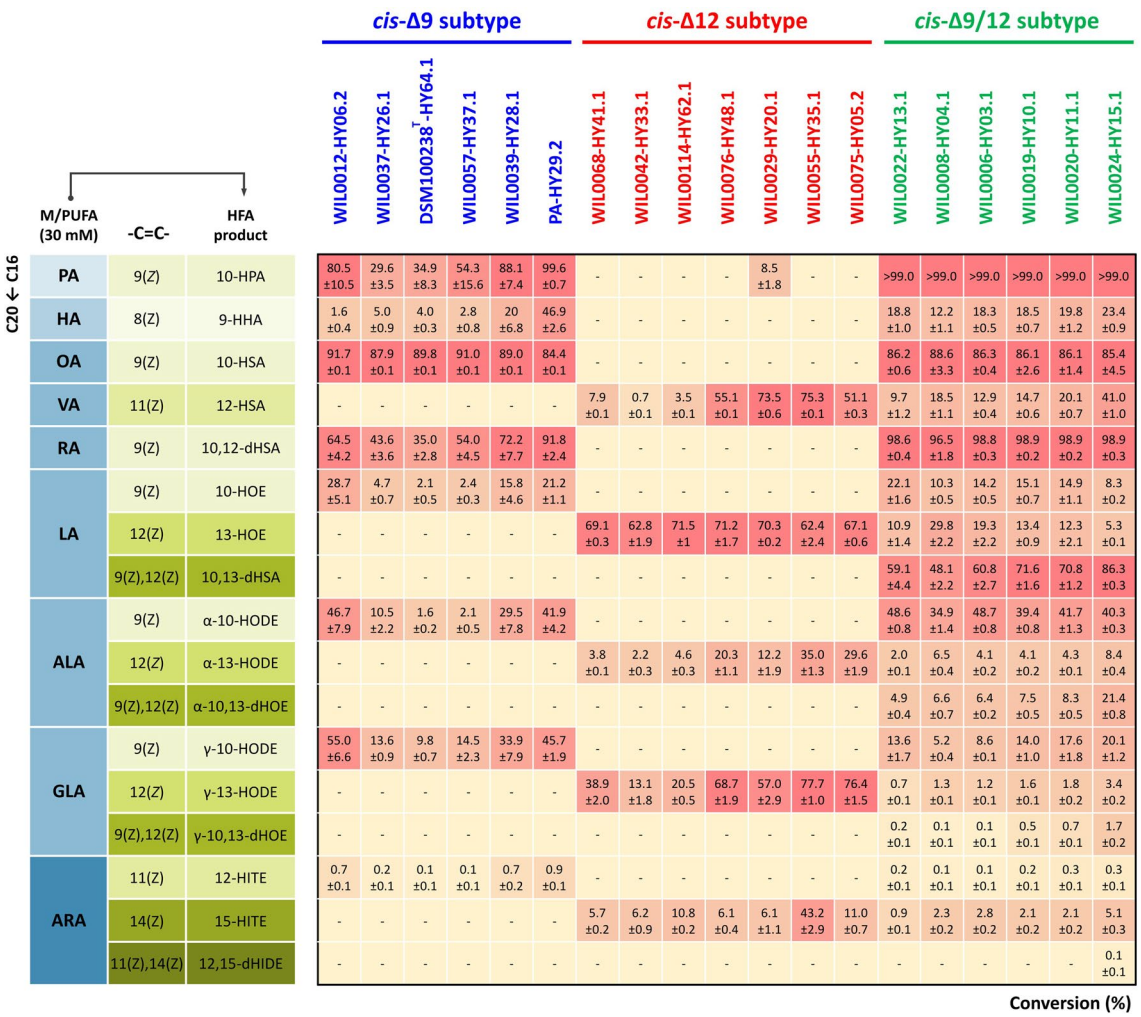


Fig. 4 Characterization of regioselectivity, substrate selectivity, and product scopes. A total of 19 FAHs representing the top performers of each subtype were selected for characterization with a panel of nine different M/PUFA substrates. The gradient from light to dark blue indicates increasing fatty acid hydrocarbon chain length. The gradient from light green to dark green indicates different *cis*-double bond positions. The heat map shows the mean percentage conversion ± standard deviation of three independent replicates (yellow, no conversion; red, 100% conversion) after 24 h of incubation at 40 °C. For fatty acid abbreviations, refer to Table 2

length with a lower degree of unsaturation. Additionally, we noted conversion of VA to 12-hydroxystearic acid (12-HSA; 0.7–75.3%) and ARA to (5*Z*,8*Z*,11*Z*)-15-hydroxyicosa-5,8,11-trienoic acid (15-HITE; 5.7–43.2%), suggesting hydroxylation of the *cis*-11 and *cis*-14 double bond of a C18 MUFA and a C20 PUFA, respectively.

All six FAHs of the *cis*-Δ9/12 subtype showed activity towards the *cis*-9 and/or *cis*-12 double bonds, converting PA, OA, RA, LA, ALA, and GLA into the respective 10-, 13-HFA and/or 10,13-di-HFA products. A higher regioselectivity for the *cis*-9 over the *cis*-12 double bond was inferred, as evidenced by a higher ratio of 10-HFA over 13-HFA intermediates in reactions involving PUFAs containing both double bonds (LA, ALA, and GLA). The accumulation of the 10-HFA intermediate resulted from more efficient hydroxylation of the *cis*-9 double bond of the PUFA substrates, coupled with less efficient hydroxylation of the *cis*-12 double bond of the generated 10-HFA intermediate to form the final 10,13-di-HFA product. Conversely, a lower level of 13-HFA intermediate corresponded to a less efficient hydroxylation of the *cis*-12 double bond of the PUFA substrates and a higher efficiency in hydroxylating the *cis*-9 double bond of the 13-HFA intermediate to form the final 10,13-di-HFA product. Substrate selectivity was generally higher and comparable for C16 and C18 MUFAs with a *cis*-9 double bond: PA (> 99.0%), OA (86.1–88.6%), and RA (96.5–98.9%). For C18 PUFAs containing both the *cis*-9 and

cis-12 double bonds, we observed a preference of FAHs of the *cis*-Δ9/12 subtype for LA (88.2–99%) over ALA (48.0–70.1%) and GLA (6.6–25.2%), based on an overall higher conversion of the substrate into the mono-HFA intermediates and the final 10,13-di-HFA product. Similarly, FAHs of the *cis*-Δ9/12 subtype hydroxylated the *cis*-8 double bond of HA (12.2–23.4%), the *cis*-11 double bond of VA (9.7–41.0%), and the *cis*-11 and *cis*-14 double bonds of ARA (1.1–5.4%), producing the corresponding 9-, 12-, 15-HFA, and 12,15-di-HFA products.

There was a distinct lack of overlap in the regioselectivity between FAHs of the *cis*-Δ9 and *cis*-Δ12 subtypes, except for the *cis*-11 double bond. In this regard, the *cis*-Δ9 subtype demonstrated hydroxylation of the *cis*-11 double bond of ARA but not VA, whereas the *cis*-Δ12 subtype exhibited the opposite behaviour. Apparently, the location of the *cis*-11 double bond as counted from the ω-end of the hydrocarbon chain (ω7 position for VA, and ω9 position for ARA) affected its hydroxylation by the two different subtypes. A closer look at the hydroxylation sites of the *cis*-Δ9 subtype revealed that the targeted *cis*-8, *cis*-9, and *cis*-11 double bonds were consistently located at the ω9 position of the hydrocarbon chain, except for the *cis*-9 double bond of PA which was at the ω7 position. Conversely, the *cis*-12 and *cis*-14 double bonds targeted by the *cis*-Δ12 subtype were located at the ω6 position, except for the *cis*-11 double bond of VA which occupied the ω7 position. This analysis suggested that the regioselectivity of the analysed FAHs could be determined by the location of the double bond from the ω-end, and proximity between the ω6 and ω7 double bonds might explain the hydroxylation of VA by the *cis*-Δ12 subtype. To avoid potential confusion with previous studies, we continued to utilize the delta-nomenclature system for describing the double bonds targeted by the different subtypes. However, in the future, it should be considered to describe regioselectivity of FAHs using the omega-system, while noting exceptions for certain M/PUFA substrates. If the omega-system is adopted, the three different subtypes would be referred to as ω9*c* (*cis*-Δ9), ω6*c* (*cis*-Δ12), and ω6/9*c* (*cis*-Δ9/12).

Nevertheless, a lack of activity on the ω7 double bond of PA by FAHs of the *cis*-Δ12 subtype indicated that the structures of the hydrocarbon chains preceding and succeeding the targeted *cis*-double bond equally influenced substrate compatibility. FAHs of the *cis*-Δ9 subtype exhibited a preference for MUFAs with a straight and saturated alkyl chain of eight carbon atoms (including the carboxyl group) preceding the targeted *cis*-double bond. Increased unsaturation or variations in the number of carbon atoms in the preceding chain significantly impacted hydroxylation by the *cis*-Δ9 subtype, as evidenced by reduced conversion of HA, GLA and ARA. Notably, the *cis*-Δ9 subtype hydroxylated PA as efficiently as OA, suggesting a preference for a straight alkyl chain of 6–8 carbon atoms succeeding the targeted *cis*-double bond. Increased unsaturation in the succeeding chain adversely affected the hydroxylation efficiency of LA, ALA, GLA, and ARA.

On the other hand, FAHs of the *cis*-Δ12 subtype preferred PUFAs with a preceding alkyl chain of eleven carbon atoms, with one conjugated *cis*-double bond ahead of the targeted *cis*-double bond, and a succeeding alkyl chain of five carbon atoms in straight and saturated configuration. Deviations from these preferred substrate structures led to decreased hydroxylation of VA, ALA, GLA, and ARA by the *cis*-Δ12 subtype. Lastly, FAHs of the *cis*-Δ9/12 subtype exhibited activities reminiscent of a blend of the *cis*-Δ9 and *cis*-12 subtypes. This subtype preferred MUFAs with straight alkyl chains of eight carbon atoms preceding and 6–8 carbon atoms succeeding the targeted *cis*-double bond, and PUFAs with two targeted *cis*-double bonds conjugated to one another, preceded by a straight alkyl chain of eight carbon atoms, and succeeded by a straight alkyl chain of five carbon atoms. The presence of additional *cis*-double bonds in the preceding or succeeding alkyl chain significantly decreased the hydroxylation of ALA, GLA, and ARA by the *cis*-Δ9/12 subtype. Considering that the presence of an additional *cis*-double bond introduces an extra bend or "kink" in the hydrocarbon chain, it is hypothesised that the drop in the hydroxylation of less preferred M/PUFA substrates could be due to a suboptimal 3D structure that is less compatible with the protein's binding site.

Next, we compared the performance of FAHs between and within the three different subtypes. In terms of MUFA substrates, FAHs of the *cis*-Δ9/12 subtype performed better than those of the *cis*-Δ9 subtype in the conversion of PA (>99% versus 29.6–99.6%) and RA (96.5–98.9% versus 35.0–91.8%), but slightly poorer than those of the *cis*-Δ12 subtype in the conversion of VA (9.7–41.0% versus 0.7–75.3%; Fig. 4). The *cis*-Δ12 subtype outperformed the *cis*-Δ9 subtype in the utilization of most PUFAs, including LA (62.4–71.5% versus 2.1–28.7%), GLA (13.1–77.7% versus 9.8–55.0%) and ARA (5.7–43.2% versus 0.1–0.9%). Conversely, the *cis*-Δ9/12 subtype excelled over the *cis*-Δ12 subtype in the conversion of LA and ALA but was less favourable in the utilization of GLA and ARA. Within FAHs of the *cis*-Δ9 subtype, PA-HY29.2 was the most versatile, showing relatively higher conversions across all eight utilizable M/PUFAs substrates. In contrast, among FAHs of the *cis*-Δ12 subtype, WIL0055-HY35.1 was relatively efficient in utilizing VA, LA, ALA, and GLA, while also demonstrating exceptional conversion of ARA to 15-HITE (43.2%). The isomers of 15-HITE, such as

(8Z,11Z,13Z)-15-hydroxyicosa-8,11,15-trienoic acid, are oxylipins with potential biological functions, including anti-carcinogenic, anti-epidermal proliferation and platelet regulation [17–20]. Additionally, WIL0029-HY20.1 of the *cis*-Δ12 subtype exhibited unique promiscuity towards the *cis*-9 double bond of PA (8.5% conversion). Lastly, WIL0024-HY15.1 was the most efficient FAH of the *cis*-9/12 subtype, showing highest conversions across most M/PUFA substrates, including HA, VA, ALA, GLA, and ARA.

The close clustering of these high-performing FAHs in the phylogenetic tree (Fig. 1b) implied a correlation between protein performance and their intrinsic sequence information. This association was particularly notable for the *cis*-Δ9 subtype, where FAHs of HFam 3 generally outperformed those of HFam 2 (no noticeable difference in recombinant expression; Fig. S2). Further investigations are warranted to elucidate the impact of protein sequence on the resulting protein structure and characteristics.

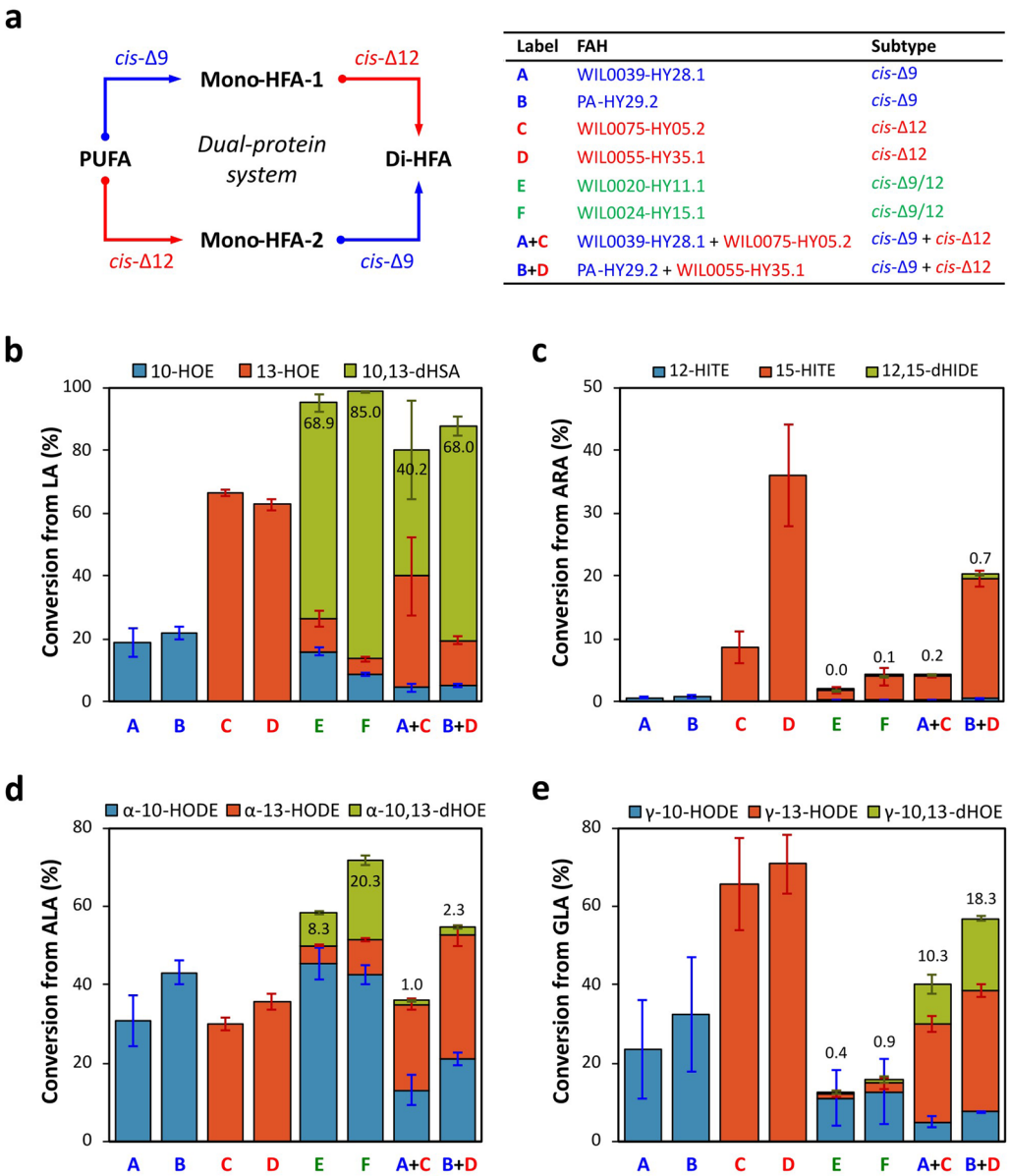
Dual-protein system for di-HFA biosynthesis

The synthesis of di-HFAs from PUFAs by the *cis*-Δ9/12 subtype was primarily hindered by a higher regioselectivity towards the *cis*-9 over the *cis*-12 double bond, as well as towards the *cis*-14 over the *cis*-11 double bond. This bias led to the accumulation of 10- or 15-HFA intermediates, which were less efficiently converted into the desired di-HFA products. To address this challenge, we proposed a dual-protein system involving a FAH of the *cis*-Δ9 subtype and a FAH of the *cis*-Δ12 subtype, working sequentially to convert PUFAs into di-HFAs (Fig. 5a). The rationale behind this strategy was to leverage the optimal regioselectivity of each subtype for specific double bond positions, thereby improving the overall conversion of PUFA substrates into di-HFA products via mono-HFA intermediates.

To test this hypothesis, we paired the top two performing FAHs of the *cis*-Δ9 and *cis*-Δ12 subtypes and co-incubated their respective induced recombinant *E. coli* cells with 30 mM of LA, ALA, GLA, and ARA substrates. The resulting fatty acid profiles were compared with those generated by the top two performing FAHs of the *cis*-Δ9/12 subtype (Fig. 5). Both *cis*-Δ9 + *cis*-Δ12 pairs formed higher levels of di-HFAs from GLA (10.3–18.3% versus 0.4–0.9%) and ARA (0.2–0.7% versus 0.1%) but not LA (40.2–68.0% versus 68.9–85.0%) and ALA (1.0–2.3% versus 8.3–20.3%). The relative levels of mono-HFA intermediates for each pair conformed with their individual selectivity for the PUFAs and their efficiency in generating the corresponding mono-HFAs. For instance, the higher selectivity for LA and greater conversion to 13-HOE by WIL0075-HY05.2 compared to 10-HOE by WIL0039-HY28.1 aligned with the relatively higher proportion of 13-HOE than 10-HOE intermediates in the dual-protein systems. It was consistently observed that FAHs of the *cis*-Δ9/12 subtype were better in the conversion of LA and ALA, whereas FAHs of the *cis*-Δ12 subtype were more efficient in the utilization of GLA and ARA.

The observed differences in the efficiency of synthesizing di-HFAs from the four PUFAs using the dual-protein systems could be attributed to variations in the efficiency of utilizing the PUFAs and their corresponding intermediates as substrates. We propose that increased production of 6(*Z*)-10,13-dihydroxyoctadec-6-enoic acid (γ-10,13-dHOE) and 12,15-dHIDE stemmed from the higher selectivity and activity of the *cis*-Δ9 + *cis*-Δ12 pair in converting GLA and ARA into mono-HFA intermediates, and the *cis*-Δ9 subtype in converting the γ-13-HODE and 15-HITE intermediates produced by the *cis*-Δ12 subtype into the corresponding di-HFA products. Both γ-13-HODE and 15-HITE lacked a *cis*-double bond succeeding the *cis*-double bond targeted by the *cis*-Δ9 subtype. The straight and saturated configuration of their succeeding alkyl chains was highly preferred by the *cis*-Δ9 subtype, resulting in more efficient conversion compared to GLA and ARA. In contrast, the γ-10-HODE and 12-HITE intermediates produced by the *cis*-Δ9 subtype were less compatible (than GLA and ARA) with the *cis*-Δ12 subtype, due to the absence of a conjugated *cis*-double bond preceding the targeted *cis*-double bonds. This differential selectivity was advantageous for the dual-protein system, as it directed the reaction first towards the *cis*-Δ12 subtype and then *cis*-Δ9 subtype, minimizing substrate competition and thereby permitting and improving the production of di-HFAs.

A similar scenario was initially theorized for the dual-protein system involving LA. This was because, in comparison to LA, the 13-HOE intermediate displayed a more favourable structure for the *cis*-Δ9 subtype, and the 10-HOE intermediate was less preferred by the *cis*-Δ12 subtype. However, this differential selectivity did not result in a higher level of 10,13-dHSA than that observed for the *cis*-Δ9/12 subtype, presumably due to a collectively lower selectivity and efficiency of the *cis*-Δ9 + *cis*-Δ12 pair than the *cis*-Δ9/12 subtype in converting LA and the mono-HFA intermediates. Nonetheless, the level of 10,13-dHSA was the highest among the di-HFAs produced from the four PUFAs using the dual-protein system. Lastly, the lower production of (15*Z*)-10,13-dihydroxyoctadec-15-enoic acid (α-10,13-dHOE) from ALA could simply be attributed to the lower selectivity and activity of the *cis*-Δ9 + *cis*-Δ12 pair in converting ALA and its mono-HFA intermediates, the structures of which were suboptimal.



Label	FAH	Subtype
A	WIL0039-HY28.1	*cis*-Δ9
B	PA-HY29.2	*cis*-Δ9
C	WIL0075-HY05.2	*cis*-Δ12
D	WIL0055-HY35.1	*cis*-Δ12
E	WIL0020-HY11.1	*cis*-Δ9/12
F	WIL0024-HY15.1	*cis*-Δ9/12
A+C	WIL0039-HY28.1 + WIL0075-HY05.2	*cis*-Δ9 + *cis*-Δ12
B+D	PA-HY29.2 + WIL0055-HY35.1	*cis*-Δ9 + *cis*-Δ12

Fig. 5 Dual-protein system for di-HFA biosynthesis. **a** The dual-protein system was constructed by pairing a top-performing FAH of the *cis*-Δ9 subtype with a top-performing FAH of the *cis*-Δ12 subtype. The FAH pair worked sequentially to convert a PUFA into mono-HFA intermediates and then into the di-HFA product. The efficiency of the dual-protein systems in generating di-HFAs from (**b**) LA, (**c**) ARA, (**d**) ALA, and (**e**) GLA was compared against those produced by the top two performing FAHs of the *cis*-Δ9/12 subtype. The bar charts show the mean percentage conversion ± standard deviation of three independent replicates after 24 h of incubation at 40 °C. The numerical values above the bars indicate the mean percentage conversion of di-HFAs. For fatty acid abbreviations, refer to Table 2

The dual-protein system demonstrated the feasibility of producing di-HFAs, or even poly-HFAs, by using multiple FAHs of different subtypes and regiospecificity. The efficiency of this system hinged significantly on the compatibility and selectivity of the FAH pair with the PUFA substrates and the HFA intermediates. Importantly, this system could enhance the production of di- or poly-HFAs from PUFA substrates less preferred by the *cis*-Δ9/12 subtype, as evidenced by the > twofold improvement in the production of γ-10,13-dHOE and 12,15-dHIDE. Future experiments are warranted to explore different FAH pairs or combinations to further enhance the level and broaden the spectrum of bioproducible di- and poly-HFAs. In this context, the *cis*-Δ12 + *cis*-Δ9/12 pair appears to be a promising combination. However, it is important to consider their potential competition for

substrates with *cis*-12 or *cis*-14 double bonds, which may compromise reaction efficiency.

Conclusions

In this study, we identified 23 FAHs using an integrated bioinformatics workflow and validated their fatty acid-hydroxylation activities via a whole-cell biotransformation assay. Further selection of 19 top-performing FAHs for characterization with nine different M/PUFA substrates shed light into their regioselectivity, substrate selectivity and product scopes. Our study proposed that FAH proteins can be classified into three distinct subtypes. The *cis*-Δ9 subtype was regioselective towards the *cis*-8, *cis*-9, and *cis*-11 (ω7 or ω9) double bonds of C16-C20 M/PUFAs, producing 9-, 10- and 12-HFAs. In contrast, the *cis*-Δ12 subtype demonstrated high regioselectivity towards the *cis*-11, *cis*-12, and *cis*-14 (ω6 or ω7) double bonds of C18 and C20 M/PUFAs, generating 12-, 13- and 15-HFAs. The *cis*-Δ9/12 subtype could hydroxylate all *cis*-double bond positions targeted by the other two subtypes, ultimately yielding di-HFAs in addition to the mono-HFA intermediates.

Our findings on the distinct requirements and preferences for specific substrate structures among the various FAH subtypes offer valuable insights for future investigations into the sequence-structure–function relationship in this domain. Such understanding paves the way for the design and optimization of FAH variants with enhanced characteristics. Indeed, we demonstrated in this study the applicability of this knowledge in designing a dual-protein system to improve di-HFA (or possibly poly-HFA) production. Specifically, through the cooperative interaction of two FAHs of the *cis*-Δ9 and *cis*-Δ12 subtypes, we improved production of two di-HFAs from PUFA substrates and mono-HFA intermediates that were less compatible with FAHs of the *cis*-Δ9/12 subtype. Importantly, one of these, 12,15-dHIDE, represented a novel di-HFA with unknown functions. The exploration of the functions of rare or novel di- and poly-HFAs would be impossible without a suitable platform that enables their feasible production. Thus, the dual-protein system developed in this study is crucial, as it offers an alternative approach for feasible production of di- and poly-HFAs.

The discovery of 23 FAHs with characteristics distinct from previously described FAHs underscores the significance of screening and characterizing unstudied FAH homologs. Importantly, half of these FAHs represent new members of the understudied *cis*-Δ12 (seven proteins) and *cis*-Δ9/12 (six proteins) subtypes. In particular, WIL0055-HY35.1 and WIL0029-HY20.1 of the *cis*-Δ12 subtype were distinctive in that the former showed remarkable synthesis of 15-HITE while the latter showed promiscuity to the *cis*-9 double bond of PA. Future study should focus on identifying the underlying key amino acid motifs or residues in these FAHs that contribute to their unique characteristics. Such insights would aid future design and development of specific FAH variants tailored towards certain applications. Finally, the availability of a FAH protein library holds considerable importance, as it enables the selection of optimal FAH proteins or combinations to facilitate commercially viable biosynthesis of HFAs.

Materials and methods

In silico gene mining and sequence similarity analysis

Homology search was performed using the *tblastn* program of BLAST + software [21], on a curated database made up of draft genomes of 139 in-house bacterial strains [14–16]. Putative FAHs were identified by alignment with the consensus protein sequence of the HMM of the MCRA protein family, using an *e*-value of less than 10^{-3} as the threshold for filtering [22]. The raw HMM of MCRA protein family was obtained from the Pfam database (https://pfam.xfam.org/) [23]. An all-versus-all, pairwise sequence comparison between the 108 putative FAHs and 55 previously described FAHs (Table S1) was performed using the online EFI-EST tool (https://efi.igb.illinois.edu/efi-est/) [24]. The resulting sequence similarity network was visualized and analysed using the organic layout of Cytoscape software (version 3.8.0) [25]. Classification into HFams was based on the 62% sequence identity threshold recommended by the Hydratase Engineering Database (https://hyed.biocatnet.de/) [11]. To infer phylogeny based on protein sequences, a multiple sequence alignment was first generated using the online MUSCLE tool [26] of EMBL-EBI (https://www.ebi.ac.uk/Tools/msa/muscle/) [27] and trimmed to remove poorly aligned or gapped positions using the "automated1" method of trimAl (version 1.3; https://vicfero.github.io/trimal/) [28]. A maximum-likelihood tree was reconstructed using IQ-TREE (version 1.6.12; http://www.iqtree.org/) [29], based on the best protein substitution model identified by the "ModelFinder" method [30] and supported with 1,000 ultrafast bootstraps [31]. Interactive Tree Of Life (version 6.8.1; https://itol.embl.de/) [32] was used for graphical visualisation of the resulting tree. The amino acid sequences of the FAHs identified and characterized in this study are provided in Table S2.

Expression plasmid construction

Genomic DNA was extracted from pure bacterial cultures using InstaGene Matrix (Bio-Rad) following the manufacturer's instructions. Specific primers were synthesised by Integrated DNA Technologies to PCR-amplify the underlying nucleotide sequences of the 23

FAHs (without start codon) from the genomic DNA. PCR reactions were conducted using Q5 High-Fidelity DNA Polymerase (New England Biolabs) with the following settings: initial denaturation at 98 °C for 60 s, followed by 35 cycles of annealing-extension at 98 °C for 10 s, 55 °C for 20 s and 72 °C for 60 s, and final extension at 72 °C for 120 s. The resulting PCR products were purified using QIAquick PCR Purification Kit (Qiagen) and cloned into the multiple cloning site of pET-15b vector (Novagen/Merck; Fig. S1a) through Gibson assembly approach. Prior to the reaction, the pET-15b vector was linearized by digestion with *BamHI* and *NdeI* restriction enzymes and purified using QIAquick PCR Purification Kit (Qiagen). Gibson assembly was performed using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs), with a vector-to-insert ratio of 1-to-2 at 50 °C for 1 h. The ligation products were transformed into the competent cells of *E. coli* strain 10-beta (New England Biolabs) and selected on lysogeny broth (LB) plates by resistance to 100 μg mL^{-1} of ampicillin. Positive transformants were confirmed by colony PCR using the primer pair T7-Pro-F (5′AGAGGATCGAGATCTCGATCCCGCG3′) and T7-Ter-R (5′CCTCTTGCGGGATATCCGGATATAG3′). Expression plasmids, pET-15b-XXX (XXX refers to the abbreviation of each FAH; Fig. S1b), were subsequently extracted from the overnight cultures of the positive transformants using QIAprep Spin Miniprep Kit (Qiagen). Lastly, DNA sequencing (1st Base Sequencing, Singapore) was conducted to validate the nucleotide sequences inserted into the vector.

Recombinant protein expression in *E. coli* strains

Expression plasmids, pET-15b-XXX, were transformed into the competent cells of *E. coli* strain Acella (Edge Biosystems) and selected on LB plates containing 100 μg mL^{-1} ampicillin to give recombinant strains Acella/pET-15b-XXX. To induce recombinant protein expression, seed cultures were first prepared by cultivating single colonies of the *E. coli* strains in 5 mL of SOC^{AMP} medium [20 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 0.5 g L^{-1} sodium chloride, 2.5 mM potassium chloride, 20 mM magnesium chloride, 20 mM magnesium sulfate, 100 μg mL^{-1} ampicillin, pH 7.0] supplemented with 20 mM glucose, at 37 °C, 220 rpm for 7 h. The obtained seed cultures were diluted to OD_{600} of 0.01 in 100 mL of autoinduction medium in 500 mL Erlenmeyer flasks and cultured aerobically with 160 rpm of agitation, first at 24 °C for 18 h and then at 14 °C for another 48 h. The autoinduction medium was prepared by supplementing the SOC^{AMP} medium with autoinduction sugars consisting of 1.5% (w/v) lactose, 0.5% (w/v) glycerol and 0.05% (w/v) glucose. The induced recombinant *E. coli* cells were harvested by centrifugation at 4000 ×*g*, 4 °C for 10 min. The collected cell pellets were washed, weighed, and resuspended to a wet cell mass of 70–90 g L^{-1} in 50 mM citrate–phosphate buffer (59.9 mM citric acid, 50 mM disodium phosphate, pH 6.0). For storage at − 80 °C, 20% (v/v) glycerol was added to the cell suspension.

To evaluate recombinant protein expression and solubility in *E. coli* strains, SDS-PAGE was conducted. Whole-cell protein extracts were prepared by lysing cells harvested from 1 mL aliquot of the induced recombinant *E. coli* cultures with 200 μL of BugBuster® Protein Extraction Reagent (Merck Millipore) following manufacturer's instructions. Soluble protein fractions were collected as supernatants of the whole-cell protein extracts after centrifugation at 21,000 ×*g* for 10 min. The remaining cell pellets were further solubilized with 8 M urea for 1 h to give insoluble protein fractions. Both soluble and insoluble protein fractions were mixed with one third volume of 4×Laemmli Sample Buffer (Bio-Rad) spiked with 1.42 M of 2-mercaptoethanol before denaturation at 95 °C for 10 min. The denatured protein samples were loaded on 4–20% Mini-PROTEAN TGX Precast Gels along with Precision Plus Protein Dual Color Standards (Bio-Rad). After electrophoresis at 120 V for 80 min with TGX buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3), the gels were stained with InstantBlue Coomassie Protein Stain (Abcam) for the visualization of protein bands.

Recombinant *E. coli* whole-cell biotransformation assays

Induced recombinant *E. coli* cells were collected by thawing the frozen cell suspensions at room temperature followed by centrifugation at 4000 ×*g*, 4 °C for 10 min to remove supernatant. After washing with 50 mM citrate–phosphate buffer (59.9 mM citric acid, 50 mM disodium phosphate, pH 6.0), the cells were resuspended in "anoxic" reaction medium, at a final wet cell mass of 35 g L^{-1}. "Anoxic" reaction medium was prepared by first adding to 50 mM citrate–phosphate buffer an appropriate amount of free fatty acid substrates and 0.25% (v/v) Tween 40, before anoxic conditioning inside a Coy anoxic chamber (atmosphere 4% hydrogen, 5% carbon dioxide and 91% nitrogen; Coy Laboratory Products) for at least 48 h. For the screening of fatty acid-hydroxylation activity, the cells were resuspended in 10 mL of "anoxic" reaction medium containing 100 g L^{-1} Wilfarin OA-7075 (Natural Oleochemicals, Malaysia), in 15 mL conical tubes. The substrate Wilfarin OA-7075 consisted of 81.0% oleic acid, 10.1% linoleic acid, 7.1% palmitic acid (C16:0) and 1.8% stearic acid (C18:0). For the characterization of substrate selectivity and product scope, cells were resuspended in 1 mL of "anoxic" reaction medium supplemented with 30 mM M/PUFA substrates, in 2 mL microcentrifuge tubes. Oleic acid (C18:1; ≥ 99% purity),

linoleic acid (C18:2; ≥ 95% purity) and arachidonic acid (C20:4; > 95% purity) were purchased from Sigma-Aldrich; vaccenic acid (C18:1; > 99% purity), ricinoleic acid (C18:1, 12-OH; > 99% purity) and α-linolenic acid (C18:3; > 99% purity) were purchased from Larodan AB; palmitoleic acid (C16:1; > 98% purity) and γ-linolenic acid (C18:3; > 98% purity) were purchased from Tokyo Chemical Industry. (8*Z*)-heptadec-8-enoic acid (C17:1) was synthesized in-house using a chemo-enzymatic approach [33, 34] and the purity was approximately 85%.

To minimize oxygen exposure, addition of cells to the "anoxic" reaction medium was conducted inside the anoxic chamber, and the tubes were sealed with parafilm before removal from the anoxic chamber. The obtained "anoxic" reaction mixtures were incubated at 40 °C with 245 rpm of agitation for 24–48 h (outside the anoxic chamber). After the reaction, samples were diluted appropriately with acetonitrile/isopropanol mixture (1-to-1 ratio) and centrifuged to settle cell debris (21,000 ×*g* for 10 min) when necessary, before LC–MS/MS analysis.

Dual-protein system

Recombinant *E. coli* cells were similarly prepared as for the whole-cell biotransformation assay. For reactions involving the interaction of two proteins, recombinant cells expressing each FAH were added to the "anoxic" reaction medium at 17.5 g L^{-1} to give a total wet cell mass of 35 g L^{-1}.

Analytical methods

Fatty acid profiles were analysed with a Vanquish LC system attached to a Q Exactive Plus Orbitrap Mass Spectrometer equipped with electrospray ionisation (Thermo Scientific). For liquid chromatography, 0.2 µL of samples were injected onto a Poroshell 120 EC-C18 (2.7 µm, 3.0 mm × 100 mm) threaded column (Agilent) at 35 °C, with a flow rate of 0.4–0.5 mL min^{-1}. Mobile phase A was 0.1% formic acid with 1 mM ammonium acetate in an acetonitrile/water mixture at 6-to-4 ratio, while mobile phase B was isopropanol. The running gradient was as follows: starting solvent mixture of 25% B; 0–2.3 min linear gradient to 50% B; 2.3–3.0 min linear gradient to 85% B; 3.0–3.5 min linear gradient to 100% B; 3.5–6.0 min isocratic with 100% B; 6.0–6.1 min linear gradient to 10% B; and 6.1–7.0 min isocratic with 10% B. For tandem mass spectrometry, samples were run in negative full scan mode followed by parallel reaction monitoring. A list of the retention times, exact masses, mass-to-charge ratios (m/z) of the precursor ions and qualifier fragment ions of M/PUFA substrates and HFA products detected and quantified in this study are provided in Table S3. Representative LC–MS/MS chromatograms and mass spectra are provided in Fig. S3-S7.

Abbreviations

FAD	Flavin adenine dinucleotide
FAH	Fatty acid hydratase
HFA	Hydroxy fatty acid
HFam	Homologous family
HMM	Hidden Markov model
LB	Lysogeny broth
LC–MS/MS	Liquid chromatography-tandem mass spectrometry
MCRA	Myosin-cross-reactive antigen
MUFA	Monounsaturated fatty acid
PUFA	Polyunsaturated fatty acid
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13068-024-02578-2.

Additional file 1.

Acknowledgements
We thank Annie Zhi Qing Sim (previously Wilmar International Limited, Singapore) for her assistance with LC-MS/MS analysis. We are grateful to Shaktheeshwari Silvaraju (Wilmar International Limited, Singapore) and Nandita Menon (previously Wilmar International Limited, Singapore) for isolating, collecting and whole-genome sequencing the bacterial strains. We further thank Prof Matthew Wook Chang, Dr Jee Loon Foo, and Dr Gazi Sakir Hossain [Synthetic Biology for Clinical and Technological Innovation (SynCTI), National University of Singapore, Singapore] for their kind contribution of the (8Z)-heptadec-8-enoic acid (C17:1) samples.

Author contributions
YC Heng: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Validation, Visualization, Writing—original draft, Writing—review & editing. GWJ Wong: Methodology, Validation, Writing—review & editing. S Kittelmann: Conceptualization, Project administration, Supervision, Writing—original draft, Writing—review & editing.

Funding
This study was funded by Wilmar International Limited, Singapore.

Availability of data and materials
Data can be made available upon request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
All authors were employed by Wilmar International Limited, Singapore. This affiliation does not alter the authors' adherence to the journal's policies on sharing data and materials.

Received: 23 July 2024 Accepted: 18 October 2024
Published online: 25 October 2024

References

1. Hira T, Ogasawara S, Yahagi A, Kamachi M, Li J, Nishimura S, et al. Novel mechanism of fatty acid sensing in enteroendocrine cells: specific structures in oxo-fatty acids produced by gut bacteria are responsible for CCK

secretion in STC-1 cells via GPR40. Mol Nutr Food Res. 2018. https://doi.org/10.1002/mnfr.201800146.
2. Furumoto H, Nanthirudjanar T, Kume T, Izumi Y, Park SB, Kitamura N, et al. 10-Oxo-trans-11-octadecenoic acid generated from linoleic acid by a gut lactic acid bacterium *Lactobacillus plantarum* is cytoprotective against oxidative stress. Toxicol Appl Pharmacol. 2016;296:1–9. https://doi.org/10.1016/j.taap.2016.02.012.
3. Ohue-Kitano R, Yasuoka Y, Goto T, Kitamura N, Park SB, Kishino S, et al. A-Linolenic acid–derived metabolites from gut lactic acid bacteria induce differentiation of anti-inflammatory M2 macrophages through G protein-coupled receptor 40. FASEB J. 2018;32(1):304–18. https://doi.org/10.1096/fj.201700273R.
4. Ikeguchi S, Izumi Y, Kitamura N, Kishino S, Ogawa J, Akaike A, et al. Inhibitory effect of the gut microbial linoleic acid metabolites, 10-oxo-trans-11-octadecenoic acid and 10-hydroxy-cis-12-octadecenoic acid, on BV-2 microglial cell activation. J Pharmacol Sci. 2018;138(1):9–15. https://doi.org/10.1016/j.jphs.2018.06.015.
5. Morito K, Shimizu R, Kitamura N, Park SB, Kishino S, Ogawa J, et al. Gut microbial metabolites of linoleic acid are metabolized by accelerated peroxisomal β-oxidation in mammalian cells. Biochim Biophys Acta Mol Cell Biol Lipids. 2019;1864(11):1619–28. https://doi.org/10.1016/j.bbalip.2019.07.010.
6. Miyamoto J, Igarashi M, Watanabe K, Karaki S, Mukouyama H, Kishino S, et al. Gut microbiota confers host resistance to obesity by metabolizing dietary polyunsaturated fatty acids. Nat Commun. 2019;10(1):1–15. https://doi.org/10.1038/s41467-019-11978-0.
7. Yamada M, Takahashi N, Matsuda Y, Sato K, Yokoji M, Sulijaya B, et al. A bacterial metabolite ameliorates periodontal pathogen-induced gingival epithelial barrier disruption via GPR40 signaling. Sci Rep. 2018;8(1):1–12. https://doi.org/10.1038/s41598-018-27408-y.
8. Miyamoto J, Mizukure T, Park SB, Kishino S, Kimura I, Hirano K, et al. A gut microbial metabolite of linoleic acid, 10-hydroxy-cis-12-octadecenoic acid, ameliorates intestinal epithelial barrier impairment partially via GPR40-MEK-ERK pathway. J Biol Chem. 2015;290(5):2902–18. https://doi.org/10.1074/jbc.M114.610733.
9. Bevers LE, Pinkse MWH, Verhaert PDEM, Hagen WR. Oleate hydratase catalyzes the hydration of a nonactivated carbon-carbon bond. J Bacteriol. 2009;191(15):5010–2. https://doi.org/10.1128/JB.00306-09.
10. Wallen LL, Benedict RG, Jackson RW. The microbiological production of 10-hydroxystearic acid from oleic acid. Arch Biochem Biophys. 1962;99(2):249–53. https://doi.org/10.1016/0003-9861(62)90006-1.
11. Schmid J, Steiner L, Fademrecht S, Pleiss J, Otte KB, Hauer B. Biocatalytic study of novel oleate hydratases. J Mol Catal B Enzym. 2016;133:S243–9. https://doi.org/10.1016/j.molcatb.2017.01.010.
12. Hirata A, Kishino S, Park SB, Takeuchi M, Kitamura N, Ogawa J. A novel unsaturated fatty acid hydratase toward C16 to C22 fatty acids from *Lactobacillus acidophilus*. J Lipid Res. 2015;56(7):1340. https://doi.org/10.1194/JLR.M059444.
13. Zhang Y, Breum NMD, Schubert S, Hashemi N, Kyhnau R, Knauf MS, et al. Semi-rational engineering of a promiscuous fatty acid hydratase for alteration of regioselectivity. ChemBioChem. 2022;23(4): e202100606. https://doi.org/10.1002/cbic.202100606.
14. Heng YC, Menon N, Chen B, Loo BZL, Wong GWJ, Lim ACH, et al. *Ligilactobacillus ubinensis* sp. nov., a novel species isolated from the wild ferment of a durian fruit (*Durio zibethinus*). Int J Syst Evol Microbiol. 2023;73(3):005733. https://doi.org/10.1099/ijsem.0.005733.
15. Heng YC, Silvaraju S, Kittelmann S. Whole-genome sequence of *Ligilactobacillus faecis* WILCCON 0062, isolated from feces of a wild boar (*Sus scrofa*). Microbiol Resour Announc. 2023. https://doi.org/10.1128/mra.00376-23.
16. Heng YC, Silvaraju S, Lee JKY, Kittelmann S. *Lactiplantibacillus brownii* sp. nov., a novel psychrotolerant species isolated from sauerkraut. Int J Syst Evol Microbiol. 2023;73(12):006194. https://doi.org/10.1099/ijsem.0.006194.
17. Yamaguchi A, van Hoorebeke C, Tourdot BE, Perry SC, Lee G, Rhoads N, et al. Fatty acids negatively regulate platelet function through formation of non-canonical 15-lipoxygenase-derived eicosanoids. Pharmacol Res Perspect. 2023;11(1): e01056. https://doi.org/10.1002/prp2.1056.
18. Ozaki N, Sakamoto N, Horikami D, Tachibana Y, Nagata N, Kobayashi K, et al. 15-Hydroxyeicosatrienoic acid induces nasal congestion by changing vascular functions in mice. Allergol Int. 2024;73(3):464–72. https://doi.org/10.1016/j.alit.2023.12.007.
19. Xi S, Pham H, Ziboh VA. 15-Hydroxyeicosatrienoic acid (15-HETrE) suppresses epidermal hyperproliferation via the modulation of nuclear transcription factor (AP-1) and apoptosis. Arch Dermatol Res. 2000;292(8):397–403. https://doi.org/10.1007/s004030000139.
20. Vang K, Ziboh VA. 15-Lipoxygenase metabolites of γ-linolenic acid/eicosapentaenoic acid suppress growth and arachidonic acid metabolism in human prostatic adenocarcinoma cells: possible implications of dietary fatty acids. Prostaglandins Leukot Essent Fat Acids. 2005;72(5):363–72. https://doi.org/10.1016/j.plefa.2005.02.002.
21. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):1–9. https://doi.org/10.1186/1471-2105-10-421.
22. Eddy S. HMMER user's guide. 2023; http://eddylab.org/software/hmmer/Userguide.pdf.
23. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: the protein families database in 2021. Nucleic Acids Res. 2021;49(D1):D412–9. https://doi.org/10.1093/nar/gkaa913.
24. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, et al. Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks. Biochim Biophys Acta Proteins Proteomics. 2015;1854(8):1019–37. https://doi.org/10.1016/j.bbapap.2015.04.015.
25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software Environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504. https://doi.org/10.1101/gr.1239303.
26. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.
27. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 2019;47(W1):W636–41. https://doi.org/10.1093/nar/gkz268.
28. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972–3. https://doi.org/10.1093/bioinformatics/btp348.
29. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74. https://doi.org/10.1093/molbev/msu300.
30. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587–9. https://doi.org/10.1038/nmeth.4285.
31. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35(2):518–22. https://doi.org/10.1093/molbev/msx281.
32. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6. https://doi.org/10.1093/nar/gkab301.
33. Travis BR, Sivakumar M, Hollist GO, Borhan B. Facile oxidation of aldehydes to acids and esters with oxone. Org Lett. 2003;5(7):1031–4. https://doi.org/10.1021/ol0340078.
34. Kaehne F, Buchhaupt M, Schrader J. A recombinant α-dioxygenase from rice to produce fatty aldehydes using *E. coli*. Appl Microbiol Biotechnol. 2011;90(3):989–95. https://doi.org/10.1007/s00253-011-3165-y.

Publisher's Note