# β-Cryptoxanthin Production in Escherichia coli by Optimization of the Cytochrome P450 CYP97H1 Activity

Thomas Lautier, Derek Smith, Lay Kien Yang, Xixian Chen, Congqiang Zhang, Gilles Truan, Nic Lindley

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

pubs.acs.org/JAFC Article

# β-Cryptoxanthin Production in *Escherichia coli* by Optimization of the Cytochrome P450 CYP97H1 Activity

*Published as part of the Journal of Agricultural and Food Chemistry virtual special issue "BIOFLAVOUR 2022 - Biotechnology of Flavours, Fragrances, and Functional Ingredients".*

Thomas Lautier,* Derek J. Smith, Lay Kien Yang, Xixian Chen, Congqiang Zhang, Gilles Truan, and Nic D. Lindley

Cite This: *J. Agric. Food Chem.* 2023, 71, 4683−4695 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Cytochromes P450, forming a superfamily of monooxygenases containing heme as a cofactor, show great versatility in substrate specificity. Metabolic engineering can take advantage of this feature to unlock novel metabolic pathways. However, the cytochromes P450 often show difficulty being expressed in a heterologous chassis. As a case study in the prokaryotic host *Escherichia coli*, the heterologous synthesis of β-cryptoxanthin was addressed. This carotenoid intermediate is difficult to produce, as its synthesis requires a monoterminal hydroxylation of β-carotene whereas most of the classic carotene hydroxylases are dihydroxylases. This study was focused on the optimization of the *in vivo* activity of CYP97H1, an original P450 β-carotene monohydroxylase. Engineering the N-terminal part of CYP97H1, identifying the matching redox partners, defining the optimal cellular background and adjusting the culture and induction conditions improved the production by 400 times compared to that of the initial strain, representing 2.7 mg/L β-cryptoxanthin and 20% of the total carotenoids produced.

**KEYWORDS:** *cytochromes P450, carotenoids, β-cryptoxanthin, metabolic engineering workflow, P450s redox toolbox*

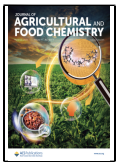


## 1. INTRODUCTION

Within the vast terpenes' family, the conjugated double bonds of carotenoids are responsible for their biological properties. In plant photosynthetic organs, carotenoids harvest and transfer energy from light to the chlorophylls through singlet excitation transfer. At saturating light intensities, carotenoids protect the photosynthetic antenna complex by absorbing excess energy through inverse transfer and releasing it by polyene vibration.[1] In nonphotosynthetic organs, carotenoids act as photoprotectors, color attractants, antioxidants, and precursors of plant hormones.[2] Even if they are not generally produced *de novo* in animals, carotenoids have essential roles such as in the composition of the rhodopsin in the retina, as antioxidants and photo protectants, or in intra- and interspecies communication (e.g., social signaling, species identification, camouflage, aposematic coloration).[2] Nonphotosynthetic microorganisms such as red yeasts or fungi also produce carotenoids involved in different nonessential functions, such as light and oxidative stress tolerances, intraspecies communication, or other ecological functions.[3] For example, the carotenoid content in *Neurospora* isolates is correlated with the latitude of their biotopes: the strains isolated from lower latitudes that receive higher UV irradiation accumulate more carotenoids.[4] More than 1100 naturally occurring carotenoids have been reported.[5] They have specific antioxidative and coloring properties, feeding useful libraries for dedicated industrial applications. For example, β-cryptoxanthin, a monohydroxylated β-carotene, is one of only six carotenoids present in human blood serum and one of the major xanthophylls that accumulate in erythrocytes.[6] Along with α-, β-, and γ- carotenes, β-cryptoxanthin is a precursor of retinoid (pro-vitamin A). In addition to its antioxidant activities, β-cryptoxanthin is associated with decreased risks of degenerative diseases and some cancers. Foods rich in β-cryptoxanthin may delay osteoporosis due to an anabolic effect on bone.[7]

Carotenoids can be produced in three ways: chemical synthesis, extraction from plants (biosourcing), or biotechnological processes. Industrial production of carotenoids has been predominantly achieved *via* chemical synthesis or extraction from plants. However, chemical synthesis results in racemic mixtures of low carotenoid bioactivity, in addition to the growing will to combine only natural compounds in food.[8] Harvesting carotenoids from plant requires large agricultural lands and is subject to geographical variations and seasonality. Moreover, the harvesting and extraction process can damage the metabolites, making the process expensive.[9] Microbial pathways can produce carotenoids fast, regardless of the seasons, at low environmental cost and with optional functionalizations, such as esterifications.[10,11] This modularity in the biotechnological approach is valuable for adding new branches to existing bioprocesses, capitalizing on existing

Received: December 21, 2022
Revised: February 19, 2023
Accepted: February 22, 2023
Published: March 8, 2023



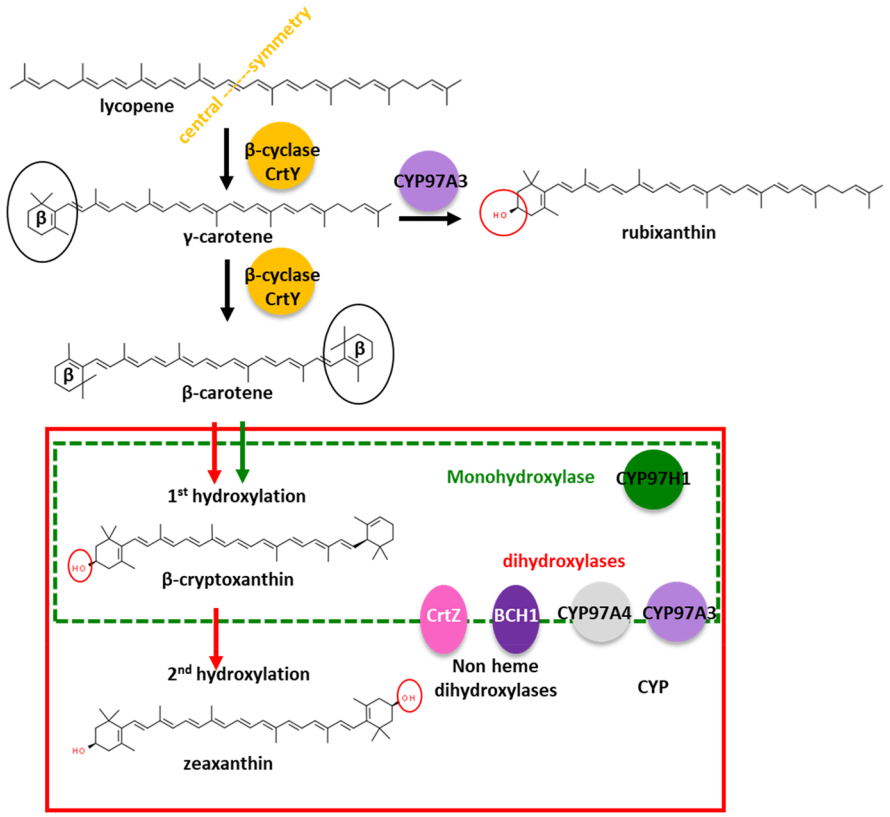



**Figure 1.** Sequential hydroxylations in the $\gamma$- and $\beta$-carotene pathways with characterized mono and dihydroxylases, belonging either to nonheme hydroxylases (CrtZ and BCH1) or to the CYP97 family (CYP97A3, CYP97A4, CYP97H1).

designs. As an example, esterifications modify the stability and the assimilation of the carotenes: natural carotenoids mainly exist as the esterified form in plants, which increases their stability, but humans can only digest the free carotenoid form. When required, biological processes allow for the addition of an esterase to produce the esterified form of the carotenoids.[12] Synthetic biology approaches have been used with success for the production of carotenoids, such as $\beta$-carotene (39.5 g/L in *Yarrowia lypolitica*). This recent breakthrough highlights the importance of protein engineering coupled with metabolic engineering in achieving industrial production.[13,14]

In addition to the end-pathway carotenoids, the molecular biodiversity libraries can be expanded with pathway intermediates, but isolating these intermediates is challenging. Chemical approaches cannot be made easily regioselective for synthesizing a dedicated intermediate metabolite. In natural sourcing, the desired intermediates accumulate at a lower level than the final carotenoids. They are also part of complex, branched metabolic pathways producing other similar intermediates, all of which complicate the product's recovery and purity. In the case of $\beta$-cryptoxantin, only one of the two identical terminal $\beta$-cycles of the precursor $\beta$-carotene is hydroxylated. Regioselective chemical syntheses for a molecule that has two identical modifiable sites are difficult to design, especially the balance of the protecting-group strategies. Although $\beta$-carotene is abundantly present in many fruits and vegetables, $\beta$-cryptoxanthin can only be found in a few species of plants, such as *Citrus unshiu*, *Citrus reticulate*, *Cucurbita maxima*, *Diospyros oleifera* Cheng, *Capsicum annuum*, and *Carica papaya*. As extraction from these species generates low yields due to the naturally low concentrations present,[15] this greatly limits the biosourcing approach. Production by biotechnology is equally challenging, as most $\beta$-carotene hydroxylases used in biotechnological processes are biterminal and produce the double hydroxylated $\beta$-carotene zeaxanthin. As such, $\beta$-cryptoxanthin is regularly cosynthesized as a minor product due to its position in the biosynthetic pathway as an intermediate product, and there have been no studies to increase production of unique oxygenated $\beta$-carotene products.[16] The $\beta$-carotene hydroxylases belong to either the nonheme hydroxylases or to the cytochromes P450 families, mainly from the CYP97 family[17] (Figure 1).

The nonheme hydroxylases are described as dihydroxylases on $\beta$-carotene and produce mainly zeaxanthin.[18] The cytochrome P450 carotene hydroxylases are mainly involved in the sequential double hydroxylation of $\alpha$-carotene (a $\beta$-carotene isomer) comprising a $\varepsilon$-cycle and a $\beta$-cycle that leads to the synthesis of lutein. In *Arabidopsis thaliana*, a sequence has been proposed for the hydroxylation of $\alpha$-carotene: CYP97A3 is first involved to hydroxylate the $\beta$-ring of $\alpha$-carotene to produce zeinoxanthin, followed by the $\varepsilon$-ring hydroxylation of zeinoxanthin by CYP97C1 to produce lutein.[19] In addition to their activity on $\alpha$-carotene, several cytochromes from the CYP97 family are known to have a minor $\beta$-carotene dihydroxylase activity, such as CYP97A3 from *A. thaliana* or CYP97A4 for *Oriza sativa*.[19,20] Among the cytochrome CYP97 family, one interesting case is CYP97H1 from the freshwater mixotrophic protist *Euglena gracilis* as it has been shown to act as a $\beta$-carotene monohydroxylase to produce $\beta$-cryptoxanthin.[21]

To accumulate $\beta$-cryptoxanthin, the substrate proximity between $\alpha$-carotene and $\beta$-carotene and the sequentiality of the hydroxylations made cytochromes P450 good candidates for the heterologous monohydroxylation of $\beta$-carotene. Hetero-


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

**Table 1. Strains and Plasmids Used in This Study**

| strain | strain background | specific plasmids | source |
|---|---|---|---|
| AB09 | BL21 (DE3) | p15A-spec-aroC-hmgS-atoB-hmgR + p15A-cam-aroB-mevK-pmk-pmd-idi + p15A-kan-aroA-IspA-CrtE-CrtB-CrtI + p15A-amp-CrtY-CrtZ2 from *P. ananatis* (strain LMG 20103) | Zhang et al., 2018[5] |
| ETL66 | MG1655 ΔaroABC | p15A-spec-aroC-hmgS-atoB-hmgR + p15A-cam-aroB-mevK-pmk-pmd-idi + p15A-kan-aroA-IspA-CrtE-CrtB-CrtI | Zhang et al., 2018[5] |
| ETL84 | BL21 (DE3) | p15A-spec-aroC-hmgS-atoB-hmgR + p15A-cam-aroB-mevK-pmk-pmd-idi + p15A-kan-aroA-IspA-CrtE-CrtB-CrtI + pTL45: p15A-amp-CrtY - T7 ferredoxin/ferredoxin reductase *S. oleracea*- T7 bov-CYP97H1 | this work |
| ETL85 | ETL66 | pTL45: p15A-amp-CrtY - T7 ferredoxin/ferredoxin reductase *S. oleracea*- T7 bov-CYP97H1 | this work |
| GroESL | BL21 (DE3) | BL21(DE3) co expressing the chaperones GroESL under T7 promoter integrated at the *adh* locus | Shukal et al., 2022[40] |
| ETL90 | BL21 adh::T7-GroESL | p15A-spec-aroC-hmgS-atoB-hmgR + p15A-cam-aroB-mevK-pmk-pmd-idi + p15A-kan-aroA-IspA-CrtE-CrtB-CrtI | this work |
| ETL99 | ETL90 | pTL45: p15A-amp-CrtY - T7 ferredoxin/ferredoxin reductase *S. oleracea*- T7 bov-CYP97H1 | this work |
| ETL100 | ETL90 | pTL43: p15A-amp-CrtY(GS)CYP97H1 - T7 ferredoxin/ferredoxin reductase *S. oleracea* | this work |
| ETL105 | ETL90 | pTL57: p15A-amp-CrtY - T7 $_{Cit}$CYP97 *C. unshiu*-rbs-CPR *C. tropicalis* | this work |
| ETL106 | ETL90 | pTL58: p15A-amp-CrtY - T7 CYP97 K15747 *C. maxima*-rbs-CPR *C. tropicalis* | this work |
| ETL119 | ETL90 | pTL63: p15A-amp-CrtY - T7 bov-CYP97H1 - T7 CPR *C. tropicalis* | this work |
| ETL125 | ETL90 | pTL104: p15A-amp-CrtY - T7 bov-CYP97H1(GS)CPR *C. tropicalis* | this work |
| ETL126 | ETL90 | pTL103: p15A-amp-CrtY- T7 bov-CYP97H1 - T7 FldA/ferredoxin reductase *E. coli* | this work |
| ETL131 | ETL90 | pTL105: p15A-amp-CrtY - T7 ferredoxin/ferredoxin *reductase S. oleracea* - T7 fl-BCH1 | this work |
| ETL135 | ETL90 | pTL62: p15A-amp-CrtY - T7 bov-CYP97H1(GS)ferredoxin -rbs-ferredoxin reductase *S. oleracea* | this work |
| ETL136 | ETL90 | pTL41: p15A-amp-CrtY - T7 bov-CYP97H1 - no redox partner | this work |
| ETL137 | ETL90 | pTL49: p15A-amp-CrtY - T7 bov-CYP97H1 - T7 ferredoxin/ferredoxin reductase *E. coli* | this work |
| ETL140 | ETL90 | pTL72: p15A-amp-CrtY - T7 ferredoxin/ferredoxin reductase *S. oleracea* - T7 tCYP97H1 | this work |
| ETL144 | ETL90 | pTL76: p15A-amp-CrtY - T7 CYP97 K15747 *C. maxima* - T7 ctCPR | this work |
| ETL150 | ETL90 | p15A-amp-CrtY | Zhang et al., 2018[5] |
| ETL152 | ETL90 | pTL87: p15A-amp-CrtY - T7 t$_{Cit}$CYP97 *C. unshiu*-rbs-CPR *C. tropicalis* | This work |
| ETL153 | ETL90 | pTL89: p15A-amp-CrtY- T7 ferredoxin/ferredoxin reductase *S. oleracea* - T7 8RP-CYP97H1 | this work |
| ETL157 | ETL90 | pTL90: p15A-amp-CrtY - T7 tCYP97 K15747 *C. maxima* - T7 ctCPR | this work |
| ETL166 | ETL90 | pTL101: p15A-amp-CrtY - T7 CYP97A4 *O. sativa* - T7 ctCPR | this work |
| ETL167 | ETL90 | pTL102: p15A-amp-CrtY -T7 ferredoxin/ferredoxin reductase *S. oleracea* - T7 tCYP97A4 *O. sativa* | this work |
| ETL168 | ETL90 | pTL100: p15A-amp-CrtY -T7-petF/ferredoxin reductase *S. oleracea* -T7 tCYP97H1 | this work |

logously expressing P450s is not simple, and the current challenge is to build an efficient workflow for functional plant cytochromes P450 expression in microorganisms.[22] In this work, a generic sequential approach was used for cytochrome P450 expression in *E. coli*, from choosing the cytochrome P450 of interest toward its optimal cellular localization and its redox environment and to determining optimal host cellular conditions. With an improvement by 400 times compared to that of the starting strain, this initial feasibility study produced 2.7 mg/L β-cryptoxanthin, which is the first quantified metabolic engineering approach targeting this metabolite. In this strain, β-cryptoxanthin represents 20% of the total carotenoids and 50% remains the precursor β-carotene, indicating that the monohydroxylase activity could still be improved.

## 2. MATERIALS AND METHODS

**Strain and Plasmid Construction.** Strains are based for the upper part of the carotene pathway on our designed astaxanthin strain[23] as a basis. The mevalonate pathway genes were cloned into p15A-*spec-hmgS-atoB-hmgR* (L2-8) (*atoB*, *hmgS* and truncated *hmgR*) and into p15A-*cam-mevK-pmk-pmd-idi* (L2-5) (*mevk*, *pmk*, *pmd*, and *idi*). The lycopene pathway genes were cloned into p15A-*kan-crtEBI-isp* (*crtEBI* and *ispA*). This last module controls lycopene cyclization into β-carotene (crtY) and its hydroxylation (various hydroxylases with their redox partners; Table 1 and Table S1). The enzymatic fusions were designed by removing the stop codon of the enzyme located in N-terminal of the fusion, adding the coding sequence for a flexible linker composed of four repetitions of the pattern GGGGS, followed by the methionine of the second enzyme of the fusion. When required, to favor the expression of genes initially in operon, their open reading frames were switched under an independent transcriptional unit, using T7 promoter variants from.[23] Codon usage for the heterologous genes was optimized for *E. coli*. Under T7 variants, expression of genes was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG). These platform strains have led to the production of various carotenoids and derived products with yields above 500 mg/L, exceeding 25g/L for some terpenoids.[9,23]

**P450 Phylogeny and Modeling.** Focusing on the CYP97 family, known to be primarily responsible for hydroxylation of β-carotene, 102 CYP97 sequences were isolated from the Uniprot database by combining the search terms "carotene & P450", a BLAST search using CYP97H1 sequence,[21] the published sequences from the CYP97 family,[20] and by a dedicated blast to identify CYP97 isoforms in the proteomes of natural producer organisms.[15] The sequences were filtered in from 400−900 residues length, representative of a P450 classic length, and were aligned for functional phylogenetic analysis. Alignment was conducted with ClustalW, and a phylogenetic tree was generated with iTOL.[24,25] The primary amino acid sequence of CYP97H1 from *E. gracilis* was modeled using Swiss-MODEL, YASARA, and AlphaFold2,[26] with PyMOL as the model viewer.

**Tube Culture of the *E. coli* Strains.** The strains were cultivated in 2 mL of 2XPY (20 g/L peptone, 10 g/L yeast extract, and 10 g/L NaCl) supplemented with 10 g/L glycerol, 75 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), and Tween 80 0.5%, as previously described.[23] The initial culture condition was 37 °C/250 rpm until $OD_{600\ nm}$ reached ∼0.6−0.8; then, the induction was triggered by adding 0.05, 0.1, 0.2, or 0.4 mM IPTG, and a production phase was conducted at 25 °C for 48 h. In the case of auto induction, 0.1, 1, 10, or 20 mM lactose were added at the start of the


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

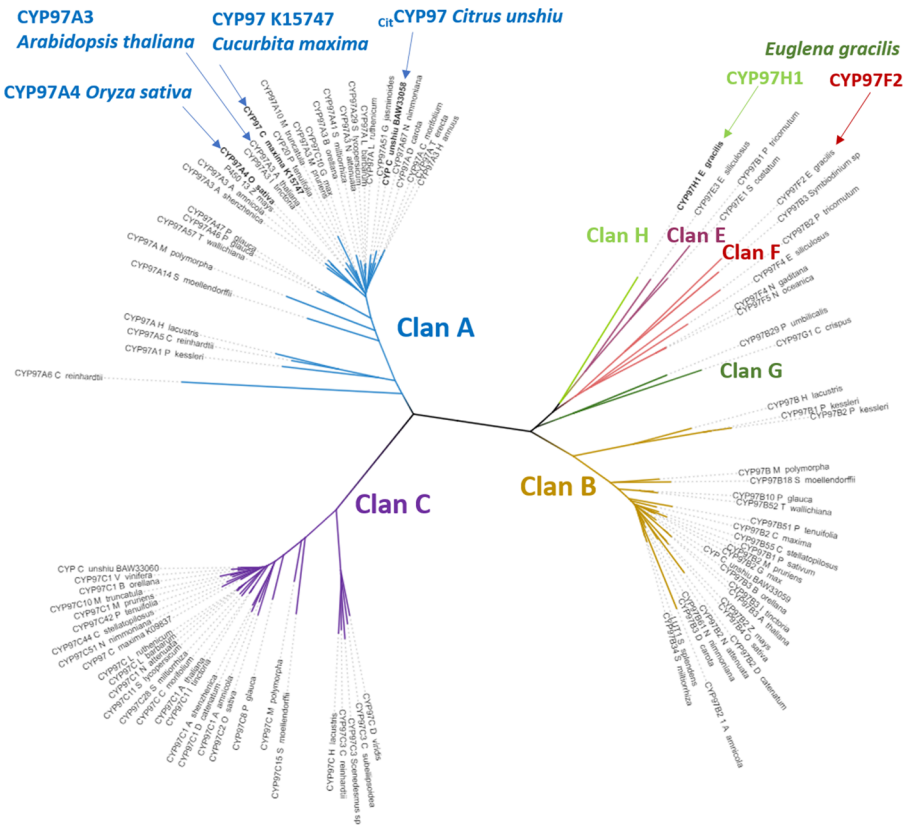



**Figure 2.** Phylogenetic tree of selected CYP97 family. The 102 CYP97 sequences from the Uniprot database, between 400 and 900 residues, are clustering in seven families: CYP97A, B, C, E, F, G, H. The CYP97 from *O. sativa* and *E. gracilis*, which are known to catalyze $\beta$-cryptoxanthin production, and *C. unshiu* and *C. maxima*, which are natural $\beta$-cryptoxanthin producers, are highlighted. The tree branches are indicating a phylogeny clustering.

culture, and the production phase was conducted at 25 °C for 48 h. To maintain the four plasmids, the antibiotics were supplemented in the culture (34 $\mu$g/mL chloramphenicol, 50 $\mu$g/mL kanamycin, 50 $\mu$g/mL spectinomycin, and 100 $\mu$g/mL ampicillin).

**Carbon Monoxide Difference Spectrum.** The method for measurement of P450 difference spectra was adapted from Johnston and Gillam, 2013.[27] After 48 h of induction, 1 mL of cells was harvested and resuspended in 500 $\mu$L of 100 mM HEPES pH 7.4, 6 mM Mg acetate, and 10 mM glucose. Two hundred microliters of cell suspension was transferred in duplicate to a microtiter plate. As in the case of an *E. coli* expression, the cytochrome P450 is already in a reduced intracellular environment; a reduction step by adding sodium dithionite was not required. CO gas was generated after chemical reaction between sulfuric acid and formic acid and neutralized by bubbling through a 0.1 M NaOH solution. Spectra were measured before and after 30 min of CO incubation, using a Biotek Synergy HT microplate reader.

**Extraction and Quantification of Carotenoids.** Briefly, 20−40 $\mu$L of microbial culture (depending on the expected content of cellular carotenoids) was collected and centrifuged. After washing the cellular pellet with water, the cells were resuspended in 20 $\mu$L of water and carotenoids were extracted by the addition of 180 $\mu$L of acetone and homogenization for 50 min at 50 °C. After centrifugation at 14 000*g* during 10 min, the supernatant was directly injected in HPLC.

**Quantification of Carotenoids.** Carotenoids were separated using an Agilent 1290 Infinity II UHPLC System and detected with a diode array detector (DAD). The analytical method was adapted from a previous protocol.[23] Briefly, 5 $\mu$L of extracted carotenoids in acetone was injected into the column Agilent ZORBAX RRHD Eclipse Plus C18 2.1 mm × 50 mm, 1.8 $\mu$m using a flow rate of 1 mL/min. The eluent gradient started with 60% methanol/water 4:1 and 40% acetonitrile for 1 min, followed by increasing acetonitrile from 40% to 80% in the following 3 min. This condition was continued for 11 min. The entire analysis finished at 15 min.

Lycopene, $\beta$-carotene, $\beta$-cryptoxanthin, rubixanthin, and zeaxanthin were purchased from CaroteNature, Münsingen, Switzerland. As the carotenoids pathway is a highly branched pathway, with several close oxygenated intermediates or isomers, lycopene, $\beta$-carotene, $\beta$-cryptoxanthin, rubixanthin, and zeaxanthin HPLC peaks were identified using a coelution standard retention time coupled to carotenoid-specific visible spectra (using the ratio between the three peaks showing the maximum of absorbance, specific for each carotene). As zeaxanthin, $\beta$-cryptoxanthin, and rubixanthin share molecular similarities, mass spectrometry fragmentation profiles were also used to confirm each peak detected by absorbance. Lycopene and zeaxanthin stock solution concentrations were quantified by dissolving a known mass in hexane. Due to standard quantity limitation, $\beta$-carotene and $\beta$-cryptoxanthin standards were dissolved in hexane, and these stock solution concentrations were calculated using the absorption coefficient $A^{1\%}_{450\ nm}$ 2500 and 2356, respectively (Sigma Chemical and reference 28). HPLC standard curves were established for the five standards using peak areas at 450 nm of serial dilution. The peak areas of each extracted compound were used to calculate the carotenoids concentrations.

Liquid chromatography−tandem mass spectrometry (LC−MS/MS) analysis was performed using an Agilent UPLC1290 coupled with a quadrupole time-of-flight (Q-TOF) system. Separation of extracts was carried out on an Agilent ZORBAX RRHD Eclipse Plus C18 (2.1 mm × 50 mm, 1.8 $\mu$m) at a flow rate of 0.5 mL/min. Both mobile phases A (water) and B (methanol) contain 0.1% formic acid. The run started initially at 90% mobile phase B for 2 min and then linearly increased to 100% B in 1.5 min and at 100% B for 8 min. The total run time is 10.5 min, followed by a postrun of 1 min.

The typical QTOF operating parameters were as follows: positive ionization mode; sheath gas nitrogen flow, 12 L/min at 295 °C;


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

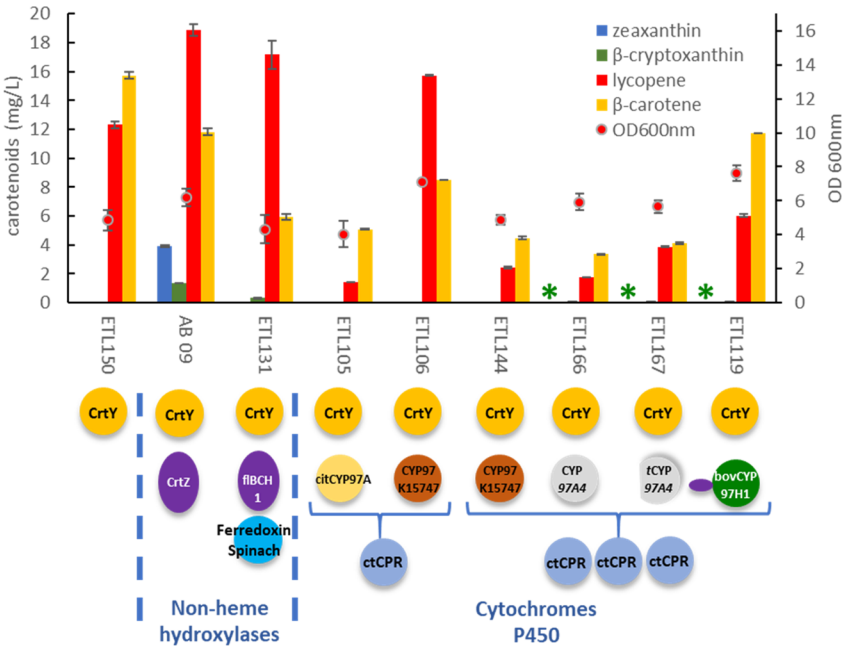



**Figure 3.** Production of carotenoids in *E. coli* expressing selected nonheme hydroxylases or cytochromes P450s. On the abscissa axis, the strains name is completed with a schematic representation of the hydroxylases and the redox partners coexpressed with the *β*-carotene cyclase CrtY from *Uncultured marine bacterium HF10_19P19*. CrtZ refers to CrtZ from *Pantoea ananatis* (strain LMG 20103), flBCH1 to the full length of BCH1 carotene hydroxylase from *A. thaliana*, "Ferredoxin Spinash" symbolizes the coexpression of the ferredoxin FER1 and the ferredoxin reductase from *S. oleracea*, $_{Cit}$CYP97A refers to the full length CYP97 of *C. unshiu*, and CYP97 K15747 describes the full length of CYP97 K15747 from *C. maxima*. ctCPR represents the cytochrome P450 reductase from *C. tropicalis*, single ctCPR symbol refers to the CPR expression in an operon, the cluster of three ctCPR symbolizes the expression of the CPR under a dedicated promoter. CYP97A4 and tCYP97A4 symbolize the full length and the truncated variant of CYP97A4 from *O. sativa*; bovCYP97H1 refers to CYP97H1 from *E. gracilis* fused to the chimeric transmembrane fragment BOV from CYP17a *B. taurus*. Strain details are available in Table 1, and protein sequences are in Table S1. On the left ordinate axis, the production is expressed in milligrams of carotenoids by liter of bacterial culture, obtained after 48 h of induction. Zeaxanthin, *β*-cryptoxanthin, lycopene, and *β*-carotene concentrations are represented, respectively, by blue, green, red, and orange bars. Green asterisks indicates a detected amount of *β*-cryptoxanthin. The right ordinate axis concerns the optical density at a wavelength of 600 nm, which is represented for each strain by a red dot. Experiments were done in triplicate, and the error bars quantify variability among triplicates.

drying gas nitrogen flow, 8 L/min at 275 °C; nebulizer pressure, 30 psi; nozzle voltage, 1.5 kV; capillary voltage, 4 kV. Lock masses in positive ion mode: purine ion at *m/z* 121.0509 and HP-921 ion at *m/z* 922.0098.

## 3. RESULTS

**3.1. Cytochromes P450 Functional Phylogenetic Analysis and Production Assay.** The CYP97 family includes key enzymes in the biosynthesis of carotenoid pigments.[29] To identify potential cytochromes P450 for *β*-cryptoxanthin production, members of the CYP97 family known to be active on the *β*-cycle of *α*-, *β*-, or *γ*-carotene such as CYP97A4 from *O. sativa*, CYP97A3 from *A. thaliana*, CYP97H1 from *E. gracilis* were chosen to construct a functional phylogenetic tree for the CYP97 family. In parallel, we explored the Uniprot database to identify CYP97 family members among organisms known to be natural high producers of *β*-cryptoxanthin[15] such as *C. unshiu*, *C. reticulate*, *C. maxima*, *Diospyros oleifera* Cheng, *Capsicum annuum*, and *Carica papaya*. Families appeared in clans A, B, C, E, F, G, and H with an over representation of the families A, B, and C and tenuous subdivisions between the clans E, F, G, and H (Figure 2).

CYP97C family members were excluded as they have been shown to be inactive on the *β*-cycle of carotene.[30] CYP97H1, thought to be a *β*-carotene monohydroxylase, did not group with other CYP members. CYP97A4, which can accumulate *β*-cryptoxanthin at certain temperatures,[20] clustered with two sequences from natural *β*-cryptoxanthin producers, the $_{Cit}$CYP97A from *C. unshiu*, known inactive on *β*-carotene,[31] and a new P450 candidate from *C. maxima*, which was worthy of assay.

Using the candidates identified in the literature and through the functional phylogenetic tree, a set of hydroxylases was expressed in a *β*-carotene producer *E. coli* strain (Figure 3).[23] In the ETL150 strain, both lycopene and *β*-carotene accumulate. At this step of the engineering, accumulation of lycopene is not problematic as the next intermediate, *β*-carotene, is available to be a substrate of the hydroxylases. Two nonheme hydroxylases, CrtZ from *P. ananatis* and BCH1 from *A. thaliana*, are both known to be zeaxanthin producers.[18,32] $_{Cit}$CYP97 from *C. unshiu* and CYP97 K15747 from *C. maxima* were selected as being the most similar to CYP97A4 from *O. sativa* and expressed in organisms known to be *β*-cryptoxanthin producers, while CYP97H1 from *E. gracilis* has been shown to be a monohydroxylase.[21] The control strain ETL150, which did not express a heterologous hydroxylase, produced around 30 mg/L carotene, with a *β*-carotene/lycopene ratio of 60:40. No *β*-cryptoxanthin or zeaxanthin were detected. In the other strains, hydroxylases and redox-partner genes were expressed on the plasmid harboring the CrtY gene (coding for *β*-carotene cyclase). Once additional genes were present, the balance for *β*-carotene production was lost and the *β*-carotene/lycopene ratio was affected, indicating that the coexpression of hydroxylases and redox partners perturbs the balance of the


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

initial synthetic metabolism. Although, the total amount of carotene remained similar in each strain, indicating that the upper part of the pathway was not affected. When the nonheme hydroxylases CrtZ and BCH1 were overexpressed (AB09 and ETL131 strains), β-cryptoxanthin was detected. The highest β-cryptoxanthin yield (1.4 mg/L) was achieved with CrtZ, a bacterial hydroxylase. Nevertheless, the CrtZ strain also accumulated zeaxanthin at 3.9 mg/L as the main hydroxylated form of β-carotene, meaning that CrtZ functions mainly as a dihydroxylase. BCH1 strain produced only β-cryptoxanthin but in a smaller amount. The metabolite profile obtained with the full length expression of BCH1 confirmed the results of the patent WO 2008/045405,[33] indicating that the full length expression of BCH1 accumulates β-cryptoxanthin, contradicting a previous study reporting BCH1 expression to produce zeaxanthin.[18] As nonheme hydroxylases are largely considered β-carotene dihydroxylases, we focus our study on the cytochrome P450 family instead.

Enzyme candidates in the cytochrome CYP97 family were expressed with the cytochrome P450 reductase from *Candida tropicalis*, ctCPR, known to supply electrons in plant P450s heterologously expressed in *E. coli*.[34] CYP97A4 from *O. sativa* expressed in a β-carotene engineered *E. coli* strain was reported to have a weak β-carotene dihydroxylase activity and to produce a mixture of β-carotene/β-cryptoxanthin/zeaxanthin.[20] Under our conditions, expressing CYP97A4 (strain ETL166) produced β-cryptoxanthin only at a low proportion, with zeaxanthin being undetectable, probably due to a low hydroxylase activity. For the heterologous expression of eukaryotic cytochrome P450 in *E. coli*, a classic approach is to delete part of the N-terminal domain that contains a peptide signal or the membrane anchor. A truncated version of CYP97A4 from *O. sativa* (strain ETL167) favored β-cryptoxanthin production. As the full length $_{Cit}$CYP97A from *C. unshiu* was inactive on β-carotene (strain ETL105), confirming previous study,[31] we assayed an N-terminal truncated version (strain ETL152) to favor the activity of the P450, but this did not produce β-cryptoxanthin. A new member of the CYP97 family from *C. maxima* was expressed for both the full-length and the truncated versions (strain ETL106 and ETL157, respectively), but both were found to not produce β-cryptoxanthin. In order to favor the expression of the P450 and the redox partner supply, the P450 and ctCPR genes were shifted from operon expression toward individual T7 promoters. Even under such conditions, neither β-cryptoxanthin nor zeaxanthin were detected (strain ETL144). These results indicated that CYP97 K15747 from *C. maxima* is not active on β-carotene when expressed in *E. coli*. The 3-D model of this P450 is found to be ∼82% identical with the CYP97A3 structure (which is mainly active on α-carotene) (Figure S1). The common P450 globular domain is then conserved in CYP97 K15747 from *C. maxima* with no additional part in the enzymatic core, which indicated the absence of introns that would need to have been cured. At the same time, the CYP97A from the green alga *Dunaliella bardawil* and $_{Cit}$CYP97A and $_{Cit}$CYP97B from *C. unshiu* are inactive on β-carotene.[31,35] Besides the usual reasons such as poor expression, poor solubility or incompatible redox enzymes, the inactivity of CYP97 K15747 from *C. maxima* on β-carotene could also be due to α-carotene being its only substrate or unknown roles of some members among the CYP97 family. The last cytochrome P450 we assayed is not from the 97A family but is the unique member of the 97H family.[21] CYP97H1 was expressed in full length, truncated in its N-terminal domain or fused with a chimeric transmembrane domain known to favor P450 expression in *E. coli*.[34] All constructs resulted in β-cryptoxanthin production (6.1 ± 0.7 μg/L in the strain ETL119). Even at an initial low production output, we decided to optimize the metabolic and protein engineering of this enzyme for two reasons: (i) CYP97H1 does not belong to the CYP97A family, for which several members are known as potential dihydroxylases under certain physiological conditions such as temperature,[20] and (ii) CYP97H1 is a unique cytochrome P450 monohydroxylase. Describing and understanding its underlying enzymatic features can result in new tools for metabolic engineering of P450s families.

**3.2. Optimization for P450 activity *in vivo*.** To optimize CYP97H1 activity *in vivo*, working hypotheses were split in two axes: the P450 heterologous host integration and its catalytic mechanism requirements. In their natural host, cytochromes P450 are involved in secondary metabolism pathways, which are spread among subcellular compartments. To be addressed into the *ad hoc* compartment, the N-terminal of cytochromes P450 often harbors a peptide signal, which is naturally cleaved to form a mature enzyme. Moreover, the involved metabolites can be hydrophobic, as the β-carotene, and are embedded in the cellular membranes. To access to their substrates, cytochrome P450 are often membrane-anchored *via* an N-terminal transmembrane helix, which is adapted to the thickness and composition of the natural host membranes. Expressed in the heterologous chassis whose membranes can be different from the natural host ones, this N-terminal extension can be deleterious and a classic strategy to improve the *in vivo* P450 productivity in *E. coli* is to modify the N-terminal membrane anchor by truncation or replacement by one favorably expressed in the host.[36] Our first working hypothesis was that tuning the N-terminal part of CYP97H1 could improve its activity in *E. coli*. Additionally to the N-terminal engineering, the expression of a eukaryotic P450 in a bacterial host requires a heterologous folding with dedicated steps, such as the heme synthesis and its insertion into the P450 catalytic site. A second approach was then to assay if the *E. coli* type of strain and the expression of chaperones could favor the CYP97H1 *in vivo* activity. Last hypothesis for an optimized cellular integration, the temperature of the culture was assayed as known to influence CYP97 behavior. Indeed with CYP97A4 from *O. sativa*, an equilibrium between mono- and dihydroxylation has been shown to be a function of temperature, with monohydroxylation increasing at a low temperature.[37] The temperature could then have an impact on CYP97H1 activity and was analyzed.

Besides optimizing the P450 expression in a heterologous cell, the requirements of its catalysis were analyzed. The P450 catalytic cycle requires two electrons given one after the other.[38] To achieve this timed electronic flux, an electron donor, such as a ferredoxin, is required. This transporter receives the electrons from a reductase, which obtains them from the NADPH cellular pool. Each cellular compartment harbors a dedicated redox partner, such as the adrenodoxin/adrenodoxin reductase in the mitochondria. A P450 targeted to a dedicated compartment is then expected to couple with the redox partner of this compartment. CYP97H1 was described to be targeted into the chloroplast of *E. gracilis*. The best redox partner for CYP97H1 could be then the ferredoxin from the chloroplast of *E. gracilis*. Its identification and expression in *E. coli* were interesting to assay. However, this matching rule is


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

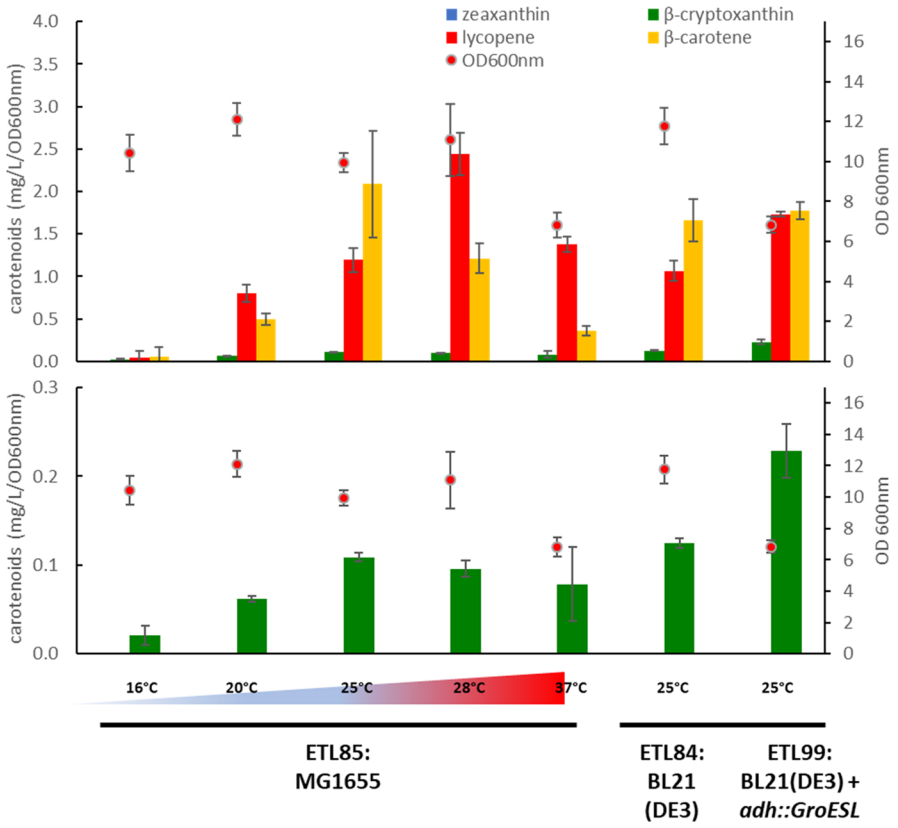



**Figure 4.** Production of carotenoids *in vivo* at different temperatures and in different strains. Bottom plot, presenting the *β*-cryptoxanthin production, is a blow-up of the top plot. The strain ETL85 expresses at 16, 20, 25, 28, and 37 °C $_{bov}$CYP97H1 (CYP97H1 from *E. gracilis* fused to the transmembrane fragment BOV from CYP17a *B. taurus*) with the ferredoxin FER1 and the ferredoxin reductase from *S. oleracea*, with the *β*-carotene cyclase in the MG1655 strain (ETL66). The strain ETL84 and ETL99 express at 25 °C the same synthetic pathways but, respectively, in BL21(DE3) and BL21(DE3) *adh::GroESL* backgrounds. Strain details are available in Table 1, and protein sequences in Table S1. On the left ordinate axis, the production is expressed in milligrams of carotenoids by liter of bacterial culture and by $OD_{600\ nm}$, obtained after 48 h of induction. Zeaxanthin, *β*-cryptoxanthin, lycopene, and *β*-carotene concentrations are represented, respectively, by blue, green, red, and orange bars. The right ordinate axis concerns the optical density at a wavelength of 600 nm, which is represented for each strain by a red dot. Experiments were done in triplicate, and the error bars quantify variability among triplicates.

not always true, especially in an heterologous context and counterexamples exist.[39] It appears then complementary to analyze the activity of CYP97H1 with a set of redox partners representative of each subcellular compartment.

The first hypothesis to test was the impact of an N-terminal engineering. CYP97H1 harbors two N-terminal transmembrane domains (TM1 residues, 15- 34, inner loop 35−93; TM2 residues, 94−116, TMHMM - 2.0 Prediction software) corresponding to a peptide signal for plastid-targeted proteins. Expressing the full length CYP97H1 in *E. coli* with the full N-terminal sequence was deleterious for the *β*-cryptoxanthin synthesis.[21] Moreover, as *β*-carotene is a hydrophobic substrate, anchoring the globular domain of CYP97H1 to the membrane could enhance its activity. In light of this, CYP97H1 was truncated of its first 116 residues and was fused to two different transmembrane domains known to increase expression of P450 in *E. coli*: the A13 fragment from CYP52A13 *C. tropicalis* and the BOV fragment from CYP17a *Bos taurus*.[36] Using pDUET plasmid expression, two chimeras were tested: the $_{bov}$CYP97H1 was coexpressed with *C. tropicalis* CPR and the $_{A13}$CYP97H1 was coexpressed with an *A. thaliana* CPR variant, in which the N-terminal domain has been replaced by the one from CPR of *C. tropicalis* for a better expression in *E. coli*. In both cases, the membraned anchored CYP97H1 produced *β*-cryptoxanthin, with a slightly better production obtained with the $_{bov}$CYP97H1 variant (data not shown). To favor accumulation of the *β*-carotene precursor, the $_{bov}$CYP97H1 chimera was coexpressed with CrtY (lycopene cyclase) on a high-copy-number plasmid with strong rbs and T7 promoter variant.[23] In this context (strain ETL85), the *β*-carotene production reached 1.5 mg/L/$OD_{600\ nm}$, allowing for the *β*-cryptoxanthin optimization assay with no limitation of the precursor.

The second hypothesis focused on the strain background, the effect of chaperones, and the cultivation temperature. Using the ferredoxin-ferredoxin reductase from *Spinacia oleracea* as the redox partner, the membrane anchored $_{bov}$CYP97H1 was expressed in the host MG1655 strain at a temperature range of 16−37 °C (strain ETL85). *β*-Carotene accumulation reached a peak at 25 °C even when lycopene content remained high. *β*-Cryptoxanthin yield was forming a bell-shaped curve from 16 to 37 °C with a maximum at 25−28 °C (Figure 4).

To optimize the cell chassis, the cell host was changed to BL21(DE3) (strain ETL84), which lead to a slightly better production of *β*-cryptoxanthin compared to the one obtained in the MG1655 strain (1.15-fold times). The production in a BL21(DE3)-GroESL context (strain ETL99) was then tested as the coexpression of the chaperones GroESL is known to facilitate the heterologous enzymes folding.[40] In this chassis


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

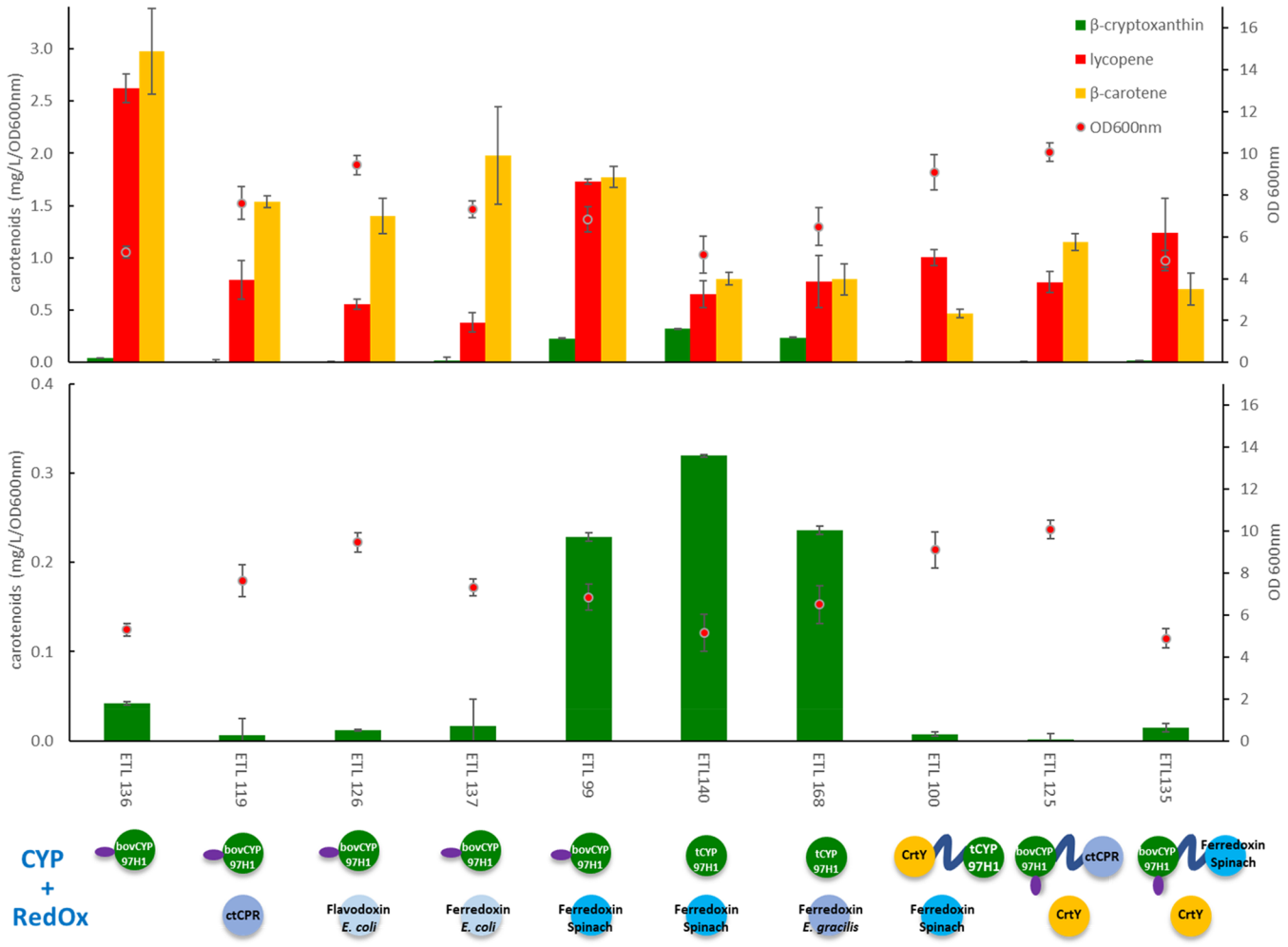



**Figure 5.** Production of carotenoids *in vivo* in strains expressing redox partners and CYP97H1 fusions with lycopene cyclase or redox partners. Bottom plot, presenting the *β*-cryptoxanthin production, is a blow-up of the top plot. On the abscissa axis, the strains name is completed with a schematic representation of the hydroxylases and the redox partners. $_{bov}$CYP97H1 refers to CYP97H1 from *E. gracilis* fused to the transmembrane fragment BOV from CYP17a *B. taurus* and tCYP97H1 refers to the truncated form of CYP97H1. ctCPR represents the cytochrome P450 reductase from *C. tropicalis*, "Flavodoxin *E. coli*" and "Ferredoxin *E. coli*" refer, respectively, to the coexpression of the flavodoxin FldA or the ferredoxin both with the ferredoxin reductase, all from *E. coli*, and "Ferredoxin Spinash" and "Ferredoxin *E. gracilis*" refer to the coexpression of the ferredoxins FER1 from *S. oleracea* or PetF from *E. gracilis* with the ferredoxin reductase from *S. oleracea*. "CrtY-tCYP97H1" refers to the C-terminal fused CrtY by a linker, composed of four repetitions of the pattern GGGGS, followed by the N-terminal of the truncated form tCYP97H1. "bovCYP97H1-ctCPR" and "bovCYP97H1-Ferredoxin Spinash" refer to the C-terminal fused bovCYP97H1 by a linker, composed of four repetitions of the pattern GGGGS, followed by the N-terminal of, respectively, ctCPR or the FER1. Strain details are available in the Table 1 and protein sequences in the Table S1. On the left ordinate axis, the production is expressed in mg of carotenoids by liter of bacterial culture and by $OD_{600\ nm}$, obtained after 48 h of induction. Zeaxanthin, *β*-cryptoxanthin, lycopene, and *β*-carotene concentrations are represented, respectively, by blue, green, red, and orange bars. The right ordinate axis concerns the optical density at a wavelength of 600 nm, which is represented for each strain by a red dot. Experiments were done in triplicate, and the error bars quantify variability among triplicates.

(strain ETL99), at 25 °C after induction, the membrane-targeting bovCYP97H1 enabled the production of 0.23 mg of *β*-cryptoxanthin/L/$OD_{600\ nm}$, which is, respectively, 1.8- and 2.1-fold times higher compared to the ETL84 BL21(DE3) and to the ETL85 MG1655 strains. Subsequent experiments were carried out at 25 °C in the BL21-GroESL chassis. The *β*-cryptoxanthin yield was then similar to the best yield achieved with the nonheme hydroxylase CrtZ (AB09 strain: 0.22 mg *β*-cryptoxanthin/L/$OD_{600\ nm}$) but without the coproduction of zeaxanthin.

The last working hypothesis to improve the activity of heterologous expressed P450s was to identify a compatible redox partner. Representative P450 redox partners belonging to each type of subcellular compartment were expressed in the BL21(DE3)-GroESL strain: (i) for endoplasmic reticulum membrane anchored P450s, microsomal cytochrome P450 reductases (from *C. tropicalis* and *A. thaliana*); (ii) for mitochondrial P450s, adrenodoxin and adrenodoxin reductase (from human); (iii) for plastid or bacterial P450s, ferredoxins and ferredoxin reductases (plant plastid from *S. oleracea*, bacterial from *E. coli*) and bacterial flavodoxin (from *E. coli*). In a tailor-made approach, the putative plastid ferredoxin PetF fragment from *E. gracilis*, the natural producer of CYP97H1, was identified in the Uniprot database but was missing the 2Fe-2S binding residues. The full sequence was identified through a genome database analysis, sharing 81% similarity with the published PetF from *Euglena viridis*. The full length


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

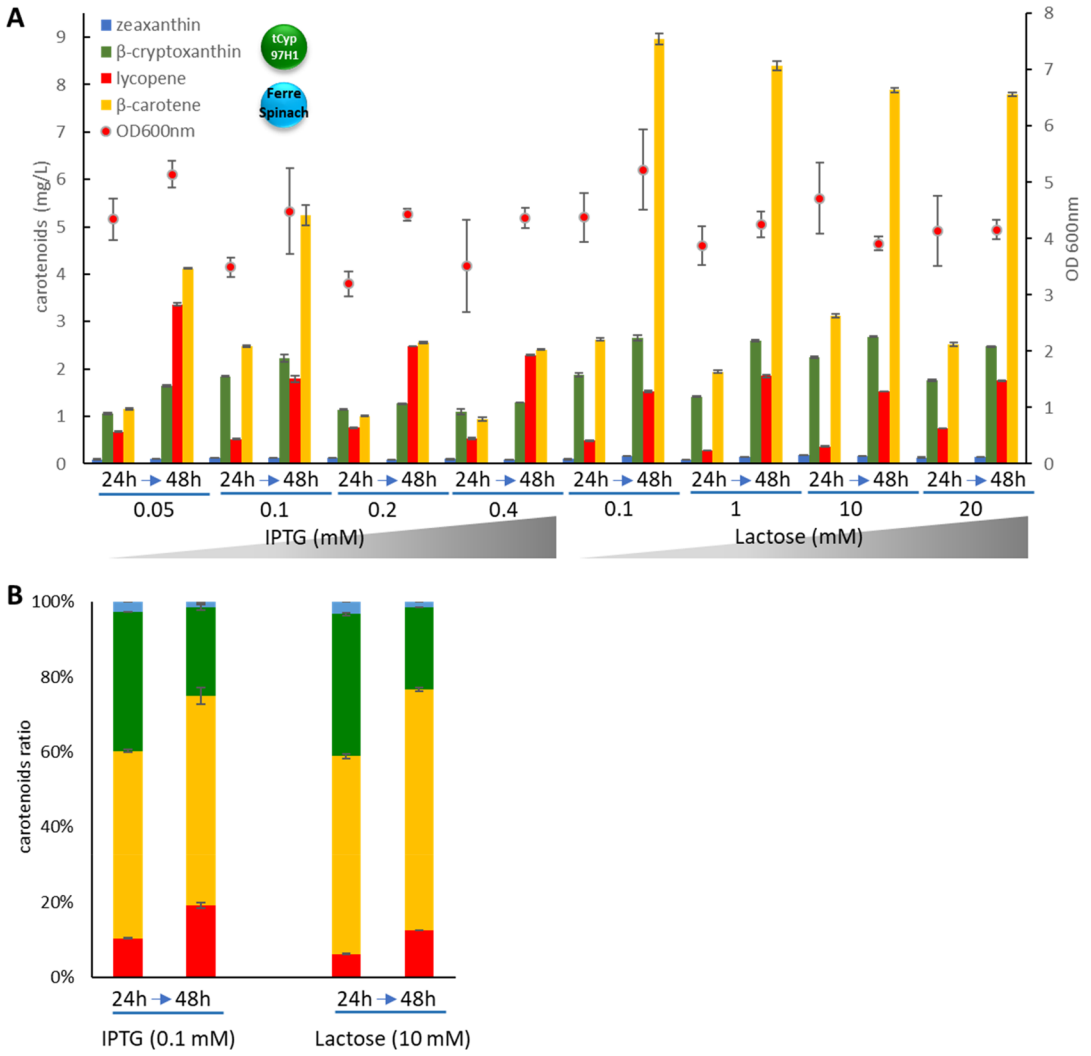



**Figure 6.** (A) Production of carotenoids *in vivo* induced by 0.05−0.4 mM IPTG or 0.1−20 mM lactose at 24 and 48 h in the strain−strain ETL140, expressing tCYP97H1 with the ferredoxin FER1 and the ferredoxin reductase from *S. oleracea* in the BL21(DE3) *adh::GroESL* β-carotene producing strain background. On the left ordinate axis, the production is expressed in milligrams of carotenoids by liter of bacterial culture. The right ordinate axis concerns the optical density at a wavelength of 600 nm, which is represented for each conditions by a red dot. (B) Carotenoids ratio at 24 and 48 h with 0.1 mM IPTG and 10 mM lactose as an inducer. Zeaxanthin, β-cryptoxanthin, lycopene, and β-carotene concentrations are represented, respectively, by blue, green, red, and orange bars. Experiments were done in triplicate, and the error bars quantify variability among triplicates.

PetF was coexpressed with the ferredoxin reductase from *S. oleracea*.

Complementary to the rational choice of the ferredoxin from *E. gracilis*, the untargeted approach using the set of redox partners was conducted using the membrane anchored $_{bov}$CYP97H1 in order to favor potential interaction with membrane bound redox partners such as the CPRs. The control strain, expressing $_{bov}$CYP97H1 without a heterologous redox partner, produced some β-cryptoxanthin (strain ETL136). In this strain, the endogenous *E. coli* redox partners, such as ferredoxin or flavodoxin, should then provide electrons to the $_{bov}$CYP97H1. In most of the combinations, a low background level of β-cryptoxanthin was detected, with no improvement in the conversion compared to the control. Overexpressing the ferredoxin or the flavodoxin from *E. coli* did not increase the β-cryptoxanthin production (Figure 5). This was surprising as ferredoxin/flavodoxin is supposed to be the endogenous redox partner whose presence explains the low background production of β-cryptoxanthin in the control strain that otherwise does not express a heterologous redox partner. Using a ferredoxin from plant plastids (from *S. oleracea*) caused a 6-fold improvement in the quantity of β-cryptoxanthin (strain ETL99). In *E. gracilis*, CYP97H1 is localized in the chloroplast, in which the redox partner is a ferredoxin. The improvement obtained by coupling CYP97H1 with a chloroplast ferredoxin highlights that the natural subcellular location of the P450 is an indication to identify a matching redox partner. Ferredoxins are not membrane-bound, and the CYP97 family members are known to be associated with the internal plastid membrane without a dedicated transmembrane domain. That is why the version of CYP97H1 with its first 116 residues truncated[21] was assayed in the β-carotene and redox partner overproducing strain and produced slightly more β-cryptoxanthin (strain ETL140: 0.32 mg of β-cryptoxanthin/L/$OD_{600\ nm}$) than the membrane-anchored variant (strain ETL99: 0.23 mg of β-cryptoxanthin/L/$OD_{600\ nm}$). This indicated that the P450/redox partner matching prevailed over the membrane anchoring in the CYP97H1 case and that a



https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

soluble form of CYP97H1 is slightly more productive when coexpressed with the ferredoxin from *S. oleracea*, potentially due to better catalytic activity, redox coupling, and/or expression efficiency. In the same view, the ferredoxin PetF from *E. gracilis* coupled with $_t$CYP97H1 and produced β-cryptoxanthin (strain ETL168) but less efficiently than with the ferredoxin from *S. oleracea*. This suggests either a lower expression of PetF or a lower affinity between PetF and a heterologous ferredoxin reductase from *S. oleracea*, highlighting the importance of a homogeneity among the redox actors.

Several examples of fused cytochrome P450 with their redox partners, named self-sufficient P450, have been reported to increase their overall activity.[41] In this view, fusion approaches were assayed with CYP97H1. The fusions built in the present work used flexible linkers that either fuse the lycopene cyclase with a truncated version of the hydroxylase or fuse the hydroxylase with a redox partner to favor the activity of the P450 itself. Both strategies impaired β-cryptoxanthin production. In the fusion CrtY-CYP97H1 (strain ETL100), the lycopene cyclase CrtY and the CYP97H1 remained active, indicating tolerance for the fusion, but the β-cryptoxanthin yield decreased. In the case of the fusions between CYP97H1 and a redox partner (CPR from *C. tropicalis* in ETL125 or ferredoxin from *S. oleracea* in ETL135), β-cryptoxanthin was detected either with the CPR or the ferredoxin fused, with the last one being more favorable probably due to its smaller size. However, in the overall redox-fused systems, the β-cryptoxanthin production collapsed. Fusing enzymes requires extensive combinations, which was not the scope of this work.[42] In subsequent experiments, we used the set of conditions that yielded optimal β-cryptoxanthin production, i.e., the BL21 (DE3) strain coexpressing the chaperones GroESL, the truncated version of CYP97H1, and the ferredoxin/ferredoxin reductase from *S. oleracea* (strain ETL140).

**3.3. Accumulation and Ratio of Carotenes.** The synthetic pathway involved 18 heterologous genes under inducible conditions that require fine induction tuning to avoid a metabolic burden. To induce the synthetic pathway in the strain ETL140, two methods were assayed: (1) a triggered induction in mid exponential growth by adding different concentrations of IPTG and (2) auto induction using lactose. Comparing these two methods showed the total carotenoid content to be higher with lactose induction, with a better conversion of lycopene into β-carotene (Figure 6A). The β-cryptoxanthin yield was slightly better in lactose induction, with the best production achieved with 10 mM lactose (2.6 mg/L β-cryptoxanthin). Traces of zeaxanthin (confirmed by mass spectrometry) were detected, suggesting a side dihydroxylase activity of CYP97H1 at an order of magnitude lower (Figure S2).

Even if the final overall production was similar at 48 h between both methods of induction, the carotenoid ratio changed from 24 to 48 h (Figure 6B). In both IPTG and lactose induction, β-cryptoxanthin represented 40% of the total carotene at 24 h. This ratio decreased to around 20% at 48 h with an accumulation of β-carotene, suggesting that the hydroxylase's bottleneck effect increased through the cellular growth. The degradation of β-cryptoxanthin in the cell, the stability of CYP97H1, or the reduction of the redox flux during growth are all hypotheses to explore. To check if the cytochrome P450 was expressed under an active form, whole cell CO spectra experiments were conducted with the strain ETL140.[27] The heme in the active pocket of a cytochrome P450 is able to bind carbon monoxide when the iron atom is in the ferrous state, its reduced form. The CO-Fe(II) versus Fe(II) differential spectra at a wavelength of 450 nm provides a sensitive means to quantify cytochrome P450 in crude mixtures. At 24 or 48 h, a P450 spectral signature was not detected even in a strain expressing only three heterologous genes with the CYP97H1 gene to favor CYP97H1 expression compared to the full system with 18 heterologous genes. Even if CYP97H1 was clearly expressed in an active form as β-cryptoxanthin was produced, these CO spectra experiments suggested that the amount of CYP97H1 was quite low. One option could be to control the expression of CYP97H1 under a specific induction system (e.g., arabinose inducible promoter) independently from the rest of the synthetic pathway, in order to over represent the mRNA of CYP97H1. In parallel, efforts to enhance proper folding can be explored, such as the addition of an N-terminal domain known to favor the expression of soluble proteins, such as a Maltose Binding Protein (MBP) tag or the 8 RP tag.[22]

## 4. DISCUSSION

In this study, an *in vivo* optimization targeting the monohydroxylase CYP97H1 from *E. gracilis* expression in *E. coli* led to a 400-fold improvement of the β-cryptoxanthin production in the strain ETL140 compared to the ETL119 starting strain. The final strain produced 2.7 mg of β-cryptoxanthin/L of microbial culture, which can be expressed as 407 mg of β-cryptoxanthin/kg of wet *E. coli* cell, considering that one unit of $OD_{600\ nm}$ corresponds to 1.7 g of cell wet weight/L of culture.[43] In plants, β-cryptoxanthin natural production can reach 34 mg/kg of butternut squash or 14 mg/kg of persimmons.[15] The higher production in the microbial approach, exempted of other oxygenated carotenoids, underlines the interest of the microbial approaches. This optimization was based on three working hypotheses generally used for P450s heterologous expression.[22]

The first angle was to test if tuning the N-terminal part of CYP97H1 could improve its activity in *E. coli*. While removing the N-terminal peptide signal has a positive effect (strain ETL140), anchoring CYP97H1 to the bacterial membrane through an N-terminal helix did not enhance the β-cryptoxanthin production (strain ETL99). The CYP97 family members are associated with plastid membranes only peripherally by a hydrophobic region.[30] As the membrane immersion depth of the P450s plays a role in catalysis, a full transmembrane domain could introduce a too rigid contact, deleterious for the activity of CYP97H1.[44] Anchoring a P450 to *E. coli* membrane is then not a guarantee of improvement and can depend on its subfamily. In a heterologous P450 expression study, both cytosolic and anchored variants should be assayed.

The second focus was to assay if the chassis strain and the culture condition affect CYP97H1 *in vivo* activity. The best chassis appeared to be BL21 strain coexpressing the chaperones GroESL, which can help to fold CYP97H1 (strain ETL99). Compared to CYP97A4 from *O. sativa*, in which the β-carotene/β-cryptoxanthin/zeaxanthin ratio shifted in favor of β-cryptoxanthin at low temperature,[37] CYP97H1 maintained the same ratio of carotenoid production independent of temperature. This focus highlights the interest of expressing cytochromes P450 with chaperones. A complementary method to favor the microbial expression of plant P450s might be to


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695

coexpress heterologous plant chaperones in the microbial chassis.[45]

The last point was identifying an efficient redox partner, essential to feed the cytochrome P450 with the required electron flux for the catalyzed oxidoreduction reaction. The redox partners of the CYP97 family members are generally thought to be plastid ferredoxins. The best redox partner for CYP97H1 could then be PetF, the ferredoxin found in the chloroplast of the same natural host *E. gracilis*. Coexpressing CYP97H1 with PetF and the ferredoxin reductase from *S. oleracea* improved β-cryptoxanthin production (strain ETL168). This was the first experimental validation that PetF is a ferredoxin. However, the highest β-cryptoxanthin yield was obtained by expressing CYP97H1 with the ferredoxin from *S. oleracea* coupled with its own ferredoxin reductase (strain ETL140). This highlights that (i) the natural redox partner is not automatically the optimal one in a heterologous context and (ii) a homogeneous relay between the electron transporter (ferredoxin), and its reducer (ferredoxin reductase) is important. In a larger view, the P450 redox partner versatility justifies then to test set of redox partners originating from each type of cellular compartment.

Taken together, these optimizations have demonstrated a pathway feasibility; however, titer is not industrially realistic for now. The evolution of the β-cryptoxanthin ratio through the time (40% of total carotene at 24 h but only 22% at 48 h in the strain ETL140) and the remaining lycopene (12% of total carotenes at 48h) indicated a pathway that currently has two controlling reactions, the lycopene cyclase and the carotene hydroxylase, which are back-to-back in the overall pathway. While flux into the pathway continues over time, the efficiency of the conversion to β-cryptoxanthin decreases. Multiple factors could explain it: β-cryptoxanthin instability, increasingly difficult thermodynamics as β-cryptoxanthin accumulates and hence slows down CYP97H1 activity, loss of activity over time due to CYP97H1 denaturation, cofactors becoming less available as growth slows and maintenance energy increases, etc.

P450s enzymes are complex with intrinsically low activity. Overexpress them to levels in which the native enzyme can perform is not realistically achievable without totally perturbing the bacterial background metabolism. Additional efforts have to be spent to redesign the P450s integration in an *E. coli* background and/or shift to a more suitable chassis in which better expression can be attained.[45−48] Heterologous metabolic circuits in classical hosts such as *E. coli*, *Saccharomyces cerevisiae*, or *Komagataella phaffii* are constantly more complex and intertwined, using system biology knowledge and technical tools gathered in the biofoundries. These evolutions are now opening roads to nonclassical chassis, such as red yeasts or filamentous fungi, potentially more favorable to express specific P450s. By applying a classical sequential strategy such as in this work and incorporating complementary approaches, such as redox partner scaffolding or synthetic tethering strategies for redox partners,[14,49] incorporating P450 enzymes into theses hosts could efficiently unlock original metabolic pathways.

Besides the choice of the cellular host, protein engineering targeting the P450s is also required to improve enzyme specific activity that will avoid crippling cells by the extent of metabolic burden. To meet the demands of P450 activity, the entire reduction chain must be adapted to a high demand for reduced cofactors, especially under growth conditions where cells can expect an increased energy demand. The redox partner then needs to be chosen including the whole redox cellular balance to build an homogeneous redox chain.[50] In addition to the redox partner engineering, the P450 heterologous workflow can be enriched targeting P450s idiosyncratic properties. For example, building a collection of N-terminal modifications will be valuable and gathering existing ones is a start. Such an N-terminal toolbox can provide the following: hydrophilic peptides (e.g.: 8RP tag); peptides to favor solubility or expression level (e.g.: MBP tag); transmembrane peptides from membrane proteins of the cellular host (e.g.: SohB$^{1-48}$ or OmpA in *E. coli*); and transmembrane peptides from CYPs belonging to the same subfamily or from different ones (e.g.: BOV, A13, EcCFS$^{1-32}$ tags).[22,36]

The use of cytochrome P450 in biotechnology cannot achieve its economic goal by using raw natural enzymatic biodiversity, which is only a starting point for such conversions, because the relevant metabolites are only produced in very small quantities in natural hosts, without strong evolutionary pressure to make these pathways efficient. Therefore, to combine the versatility of P450 substrates with efficient catalysis, iterative generic tools, such as a toolbox of redox partners tailored to the host chassis, need to be assembled and developed using metabolic and protein engineering.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

All data supporting the findings of this study are available in the article, Supporting Information, or upon request from the corresponding author.

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jafc.2c08970.

Table of protein sequences of the redox partners and the cytochromes P450 variants and figures of AlphaFold-modelled structure of CYP97 K15747 and the mass spectrometry analysis of β-cryptoxanthin and zeaxanthin in ETL140 (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

**Thomas Lautier** − *Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 138669, Singapore; TBI, Université de Toulouse, CNRS, INRAE, INSA, 31077 Toulouse, France; CNRS@CREATE, 138602, Singapore;* orcid.org/0000-0002-5465-1963; Phone: +33(0) 567048813; Email: thomas.lautier@cnrs.fr

### Authors

**Derek J. Smith** − *Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 138669, Singapore;* orcid.org/0000-0001-5103-2107

**Lay Kien Yang** − *Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 138669, Singapore*

**Xixian Chen** − *Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 138669, Singapore*

**Congqiang Zhang** − *Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 138669, Singapore*
**Gilles Truan** − *TBI, Université de Toulouse, CNRS, INRAE, INSA, 31077 Toulouse, France*
**Nic D. Lindley** − *Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 138669, Singapore; TBI, Université de Toulouse, CNRS, INRAE, INSA, 31077 Toulouse, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jafc.2c08970

### Author Contributions

The experimental work was conducted by T.L. and L.K.Y. Conceptualization: T.L., D.J.S., G.T., and N.D.L. Writing (original draft preparation): T.L. Writing (review and editing): T.L., X.C., C.Z., and N.D.L. Supervision: G.T. and N.D.L. Funding acquisition: T.L., C.Z., N.D.L., and G.T. All authors have read and approved the published version of the manuscript.

### Funding

In the frame of the France-Singapore International Research Program in Synthetic Biology for a Bioinspired Economy, this research was funded by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) EcoCTs project and by the Biotrans IAF-PP (HBMS Domain): H17/01/a0/006. The 2019 AME Young Investigator Research Grant (YIRG) A2084c0064 awarded to C.Z. supported the laboratory enzymes and carotenoid standards for this project. The analytical work was funded by the Singapore Integrative Biosystems and Engineering Research (SIBER) Grant.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors would like to thank Yoganathan Kanagasundaram for providing analytical support throughout this study, CNRS@CREATE, the Scientific Editing Team in A*STAR and Sharon Cameli for administrative supports.

## ■ REFERENCES

(1) Sandmann, G. Diversity and origin of carotenoid biosynthesis: its history of coevolution towards plant photosynthesis. *New Phytologist* **2021**, *232*, 479−493.
(2) Maoka, T. Carotenoids as natural functional pigments. *J. Nat. Med.* **2020**, *74*, 1−16.
(3) Avalos, J.; Carmen Limón, M. Biological roles of fungal carotenoids. *Curr. Genet* **2015**, *61*, 309−324.
(4) Luque, E. M.; Gutiérrez, G.; Navarro-Sampedro, L.; Olmedo, M.; Rodríguez-Romero, J.; Ruger-Herreros, C.; Tagua, V. G.; Corrochano, L. M. A Relationship between Carotenoid Accumulation and the Distribution of Species of the Fungus Neurospora in Spain. *PLoS One* **2012**, *7*, e33658.
(5) Zhang, C. Biosynthesis of Carotenoids and Apocarotenoids by Microorganisms and Their Industrial Potential. In *Progress in Carotenoid Research*; Zepka, L. Q., Jacob-Lopes, E., Rosso, V. V. D., Eds.; InTech, 2018.
(6) Nishino, A.; Ichihara, T.; Takaha, T.; Kuriki, T.; Nihei, H.; Kawamoto, K.; Yasui, H.; Maoka, T. Accumulation of Paprika Carotenoids in Human Plasma and Erythrocytes. *Journal of Oleo Science* **2015**, *64*, 1135−1142.
(7) Caballero, B. *Encyclopedia of human nutrition*; Academic Press, 1999.
(8) Snell, T. W.; Carberry, J. Astaxanthin Bioactivity Is Determined by Stereoisomer Composition and Extraction Method. *Nutrients* **2022**, *14*, 1522.
(9) Zhang, C.; Chen, X.; Too, H.-P. Production of carotenoids and apocarotenoids. US 0367928 A1, 2019.
(10) Igreja, W. S.; Maia, F. d. A.; Lopes, A. S.; Chisté, R. C. Biotechnological Production of Carotenoids Using Low Cost-Substrates Is Influenced by Cultivation Parameters: A Review. *Int. J. Mol. Sci.* **2021**, *22*, 8819.
(11) Chen, X.; Lim, X.; Bouin, A.; Lautier, T.; Zhang, C. High-level de novo biosynthesis of glycosylated zeaxanthin and astaxanthin in Escherichia coli. *Bioresour. Bioprocess.* **2021**, *8*, 67.
(12) Pérez-Gálvez, A.; Mínguez-Mosquera, M. I. Esterification of xanthophylls and its effect on chemical behavior and bioavailability of carotenoids in the human. *Nutr. Res. (N.Y.)* **2005**, *25*, 631−640.
(13) Ma, Y.; Liu, N.; Greisen, P.; Li, J.; Qiao, K.; Huang, S.; Stephanopoulos, G. Removal of lycopene substrate inhibition enables high carotenoid productivity in Yarrowia lipolytica. *Nat. Commun.* **2022**, *13*, 572.
(14) Park, S. Y.; Eun, H.; Lee, M. H.; Lee, S. Y. Metabolic engineering of Escherichia coli with electron channelling for the production of natural products. *Nat. Catal* **2022**, *5*, 726−737.
(15) Burri, B. J.; La Frano, M. R.; Zhu, C. Absorption, metabolism, and functions of β-cryptoxanthin. *Nutr Rev.* **2016**, *74*, 69−82.
(16) Niu, F.-X.; Lu, Q.; Bu, Y.-F.; Liu, J.-Z. Metabolic engineering for the microbial production of isoprenoids: Carotenoids and isoprenoid-based biofuels. *Synth Syst. Biotechnol* **2017**, *2*, 167−175.
(17) Kim, J.; Smith, J. J.; Tian, L.; Dellapenna, D. The evolution and function of carotenoid hydroxylases in Arabidopsis. *Plant Cell Physiol* **2009**, *50*, 463−479.
(18) Sun, Z.; Gantt, E.; Cunningham, F. X. Cloning and Functional Analysis of the β-Carotene Hydroxylase of Arabidopsis thaliana. *J. Biol. Chem.* **1996**, *271*, 24349−24352.
(19) Kim, J.; DellaPenna, D. Defining the primary route for lutein synthesis in plants: The role of Arabidopsis carotenoid β-ring hydroxylase CYP97A3. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 3474−3479.
(20) Quinlan, R. F.; Jaradat, T. T.; Wurtzel, E. T. Escherichia coli as a platform for functional expression of plant P450 carotene hydroxylases. *Arch. Biochem. Biophys.* **2007**, *458*, 146−157.
(21) Tamaki, S.; Kato, S.; Shinomura, T.; Ishikawa, T.; Imaishi, H. Physiological role of β-carotene monohydroxylase (CYP97H1) in carotenoid biosynthesis in Euglena gracilis. *Plant Sci.* **2019**, *278*, 80−87.
(22) Zhou, A.; Zhou, K.; Li, Y. Rational Design Strategies for Functional Reconstitution of Plant Cytochrome P450s in Microbial Systems. *Curr. Opin Plant Biol.* **2021**, *60*, 102005.
(23) Zhang, C.; Seow, V. Y.; Chen, X.; Too, H.-P. Multidimensional heuristic process for high-yield production of astaxanthin and fragrance molecules in Escherichia coli. *Nat. Commun.* **2018**, *9*, 1858.
(24) Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; Lopez, R. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636−W641.
(25) Ciccarelli, F. D.; Doerks, T.; von Mering, C.; Creevey, C. J.; Snel, B.; Bork, P. Toward automatic reconstruction of a highly resolved tree of life. *Science* **2006**, *311*, 1283−1287.
(26) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.

(27) Johnston, W. A.; Gillam, E. M. J. Measurement of P450 difference spectra using intact cells. *Methods Mol. Biol.* **2013**, *987*, 189−204.
(28) De Ritter, E.; Purcell, A. E. 10 - CAROTENOID ANALYTICAL METHODS. In *Carotenoids as Colorants and Vitamin A Precursors*; Bauernfeind, J. C., Eds.; Elsevier, 1981; pp 815−923.
(29) Hansen, C. C.; Nelson, D. R.; Møller, B. L.; Werck-Reichhart, D. Plant cytochrome P450 plasticity and evolution. *Molecular Plant* **2021**, *14*, 1244−1265.
(30) Quinlan, R. F.; Shumskaya, M.; Bradbury, L. M. T.; Beltrán, J.; Ma, C.; Kennelly, E. J.; Wurtzel, E. T. Synergistic Interactions between Carotene Ring Hydroxylases Drive Lutein Formation in Plant Carotenoid Biosynthesis. *Plant Physiology* **2012**, *160*, 204−214.
(31) Ma, G.; Zhang, L.; Yungyuen, W.; Tsukamoto, I.; Iijima, N.; Oikawa, M.; Yamawaki, K.; Yahata, M.; Kato, M. Expression and functional analysis of citrus carotene hydroxylases: unravelling the xanthophyll biosynthesis in citrus fruits. *BMC Plant Biol.* **2016**, *16*, 148.
(32) Choi, S.-K.; Matsuda, S.; Hoshino, T.; Peng, X.; Misawa, N. Characterization of bacterial beta-carotene 3,3′-hydroxylases, CrtZ, and P450 in astaxanthin biosynthetic pathway and adonirubin production by gene combination in Escherichia coli. *Appl. Microbiol. Biotechnol.* **2006**, *72*, 1238−1246.
(33) Louie, M.; Fuerst, E.-J. Biosynthesis of Beta-Cryptoxanthin. WO/2008/045405, 2008.
(34) Chang, M. C. Y.; Eachus, R. A.; Trieu, W.; Ro, D.-K.; Keasling, J. D. Engineering Escherichia coli for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol.* **2007**, *3*, 274−277.
(35) Liang, M.-H.; Xie, H.; Chen, H.-H.; Liang, Z.-C.; Jiang, J.-G. Functional Identification of Two Types of Carotene Hydroxylases from the Green Alga Dunaliella bardawil Rich in Lutein. *ACS Synth. Biol.* **2020**, *9*, 1246−1253.
(36) Chang, M. C. Y.; Eachus, R. A.; Trieu, W.; Ro, D.-K.; Keasling, J. D. Engineering Escherichia coli for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol.* **2007**, *3*, 274−277.
(37) Quinlan, R. F.; Jaradat, T. T.; Wurtzel, E. T. Escherichia coli as a platform for functional expression of plant P450 carotene hydroxylases. *Arch. Biochem. Biophys.* **2007**, *458*, 146−157.
(38) Dubey, K. D.; Shaik, S. Cytochrome P450—The Wonderful Nanomachine Revealed through Dynamic Simulations of the Catalytic Cycle. *Acc. Chem. Res.* **2019**, *52*, 389−399.
(39) Sintupachee, S.; Ngamrojanavanich, N.; Sitthithaworn, W.; De-Eknamkul, W. Molecular cloning, bacterial expression and functional characterisation of cytochrome P450 monooxygenase, CYP97C27, and NADPH-cytochrome P450 reductase, CPR I, from Croton stellatopilosus Ohba. *Plant Science* **2014**, *229*, 131−141.
(40) Shukal, S.; Lim, X. H.; Zhang, C.; Chen, X. Metabolic engineering of Escherichia coli BL21 strain using simplified CRISPR-Cas9 and asymmetric homology arms recombineering. *Microb. Cell Fact.* **2022**, *21*, 19.
(41) Correddu, D.; Di Nardo, G.; Gilardi, G. Self-Sufficient Class VII Cytochromes P450: From Full-Length Structure to Synthetic Biology Applications. *Trends Biotechnol.* **2021**, *39*, 1184−1207.
(42) Rabeharindranto, H.; Castaño-Cerezo, S.; Lautier, T.; Garcia-Alles, L. F.; Treitz, C.; Tholey, A.; Truan, G. Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in S. cerevisiae. *Metab Eng. Commun.* **2019**, *8*, e00086.
(43) Glazyrina, J.; Materne, E.-M.; Dreher, T.; Storm, D.; Junne, S.; Adams, T.; Greller, G.; Neubauer, P. High cell density cultivation and recombinant protein production with Escherichia coli in a rocking-motion-type bioreactor. *Microb. Cell Fact.* **2010**, *9*, 42.
(44) Mukherjee, G.; Nandekar, P. P.; Wade, R. C. An electron transfer competent structural ensemble of membrane-bound cytochrome P450 1A1 and cytochrome P450 oxidoreductase. *Commun. Biol.* **2021**, *4*, 1−13.
(45) Hausjell, J.; Halbwirth, H.; Spadiut, O. Recombinant production of eukaryotic cytochrome P450s in microbial cell factories. *Biosci. Rep.* **2018**, *38* (2), BSR20171290.
(46) Hu, B.; Zhao, X.; Wang, E.; Zhou, J.; Li, J.; Chen, J.; Du, G. Efficient heterologous expression of cytochrome P450 enzymes in microorganisms for the biosynthesis of natural products. *Crit. Rev. Biotechnol.* **2023**, *43*, 227.
(47) Zelasko, S.; Palaria, A.; Das, A. Optimizations to achieve high-level expression of cytochrome P450 proteins using Escherichia coli expression systems. *Protein Expr Purif* **2013**, *92*, 77−87.
(48) Li, Z.; Jiang, Y.; Guengerich, F. P.; Ma, L.; Li, S.; Zhang, W. Engineering cytochrome P450 enzyme systems for biomedical and biotechnological applications. *J. Biol. Chem.* **2020**, *295*, 833−849.
(49) Haslinger, K.; Prather, K. L. J. Heterologous caffeic acid biosynthesis in Escherichia coli is affected by choice of tyrosine ammonia lyase and redox partners for bacterial Cytochrome P450. *Microb. Cell Fact.* **2020**, *19* (1), 26.
(50) Nowrouzi, B.; Rios-Solis, L. Redox metabolism for improving whole-cell P450-catalysed terpenoid biosynthesis. *Critical Reviews in Biotechnology* **2022**, *42*, 1213−1237.


https://doi.org/10.1021/acs.jafc.2c08970
J. Agric. Food Chem. 2023, 71, 4683−4695