First Betalain-Producing Bacteria Break the Exclusive Presence of the Pigments in the Plant Kingdom

Luis Eduardo Contreras-Llano,a,b* M. Alejandra Guerrero-Rubio,a José Daniel Lozada-Ramírez,b Francisco García-Carmona,a Fernando Gandía-Herreroa

aDepartment of Biochemistry and Molecular Biology A, Faculty of Biology, Regional Campus of International Excellence, Campus Mare Nostrum, University of Murcia, Murcia, Spain
bDepartment of Health Sciences, Universidad de las Americas Puebla, Puebla, Mexico

ABSTRACT The biosynthesis of antioxidant pigments, namely, betalains, was believed to be restricted to Caryophyllales plants. This paper changes this paradigm, and enzyme mining from bacterial hosts promoted the discovery of bacterial cultures producing betalains. The spectrum of possible sources of betalain pigments in nature is broadened by our description of the first betalain-forming bacterium, Gluconacetobacter diazotrophicus. The enzyme-specific step is the extradiol cleavage of the precursor amino acid L-dihydroxyphenylalanine (L-DOPA) to form the structural unit betalamic acid. Molecular and functional work conducted led to the characterization of a novel dioxygenase, a polypeptide of 17.8 kDa with a \( K_m \) of 1.36 mM, with higher activity and affinity than those of its plant counterparts. Its superior activity allowed the first experimental characterization of the early steps in the biosynthesis of betalains by fully characterizing the presence and time evolution of 2,3- and 4,5-seco-DOPA intermediates. Furthermore, spontaneous chemical reactions are characterized and incorporated into a comprehensive enzymatic-chemical mechanism that yields the final pigments.

IMPORTANCE Several studies have demonstrated the health-promoting effects of betalains due to their high antioxidant capacity and their positive effect on the dose-dependent inhibition of cancer cells and their proliferation. To date, betalains were restricted to plants of the order Caryophyllales and some species of fungi, but the present study reveals the first betalain-producing bacterium, as well as the first steps in the formation of pigments. This finding demonstrates that betalain biosynthesis can be expanded to prokaryotes.

KEYWORDS betalains, betalamic acid, dioxygenase, enzyme mining, pigments

Betalains are pigments classified into two groups, the violet betacyanins and the yellow betaxanthins, which in addition present green fluorescence (1). The combination of both types results in the wealth of colors present in plants of the order Caryophyllales (2). Betalains are molecules with a strong antiradical capacity (3). The betalamic acid resonance system is responsible for this activity, which is modulated by the condensed molecule in each individual pigment (4). Thus, betalamic acid is not only the structural backbone of betalains but also their bioactive unit. Various studies with multiple cell lines have revealed that betalains are active against the proliferation of cancer cells (5, 6). In addition, reduction of induced tumors in vivo has been reported in mice when they were orally administered betalain pigments (7). In humans, betalain-rich extracts promoted an anti-inflammatory response (8). Their health-promoting effect has also been reported in the animal model Caenorhabditis elegans, where betalains reduce oxidative stress in vivo and increase life span (9). Thus, betalains are considered phytochemicals of nutritional value with high bioactive potential (10).
biosynthetic pathway of betalains implies the formation of betalamic acid by the
enzyme 4,5-dihydroxyphenylalanine (DOPA)-extradiol-dioxygenase (4,5-DODA) and its
further condensation with amino acids and amines (11). 4,5-DODA catalyzes the ring
opening oxidation of the molecule L-3,4-dihydroxyphenylalanine (L-DOPA) to form the
intermediate 4,5-seco-DOPA, which cyclizes spontaneously to betalamic acid (2). Analogues
of betalains are present in the fungi Amanita (12) and Hygrocybe (13), where
betalain-related pigments exist derived from muscaflavin, a betalamic acid isomer. No
evidence that bacteria may synthesize betalains exists in the literature, but our search
for novel biological systems and enzyme mining from nonnative hosts able to catalyze
this reaction led to establishing bacterial cultures of microorganisms and supplement-
ing them with l-DOPA as a precursor. This paper describes the cloning, expression,
purification, and molecular and functional characterization of the betalamic acid form-
ing DODA-extradiol-dioxygenase from Gluconacetobacter diazotrophicus, described here
to be the first bacterium producing betalains in culture. The kinetics of betalamic acid,
muscaflavin, and dopaxanthin formation from L-DOPA were characterized in real time
using kinetic monitoring of the reaction medium by high-performance liquid chroma-
tography (HPLC) analysis, which for the first time allows the visualization and full
characterization of 2,3- and 4,5-seco-DOPA intermediates and their formation.

RESULTS AND DISCUSSION

Gluconacetobacter diazotrophicus cultures produce betalamic acid. Gluconacet-
obacter diazotrophicus is a proteobacterium first described in roots and stems of
sugarcane (14) with no evident relationship with plants of the order Caryophyllales, but
its cultures supplemented with L-DOPA (7.6 mM) showed yellow coloration. HPLC
analysis of yellow G. diazotrophicus cultures showed the presence of a peak with a
retention time (Rt) of 13.67 min and with an exact mass detected by HPLC-electrospray
ionization-time of flight mass spectrometry (ESI-TOF MS) of 391.1144 m/z. This mass
corresponded to the pigment dopaxanthin, the DOPA-derived betaxanthin first de-
scribed in Glottiphyllum longum flowers (15), and its identity was corroborated using a
real dopaxanthin standard (Fig. 1). This positive result was further investigated by
harvesting G. diazotrophicus cells and resuspending them in water supplemented with
increasing concentrations of L-DOPA until its solubility limit was reached at 7.6 mM.
After 24 h, this medium showed yellow coloration, and the presence of dopaxanthin
was confirmed by HPLC-ESI-TOF MS. The exact mass of 391.1144 m/z, corresponding to
that of dopaxanthin, was detected in all the samples, and its accumulation was higher
as the concentration of L-DOPA increased (see Fig. S1 in the supplemental material). The
presence of L-DOPA stimulated the production of dopaxanthin, but small amounts of
dopaxanthin were also detected in the absence of added L-DOPA. Therefore, Gluconac-
etobacter diazotrophicus produces the betalain dopaxanthin under physiological con-
ditions in the absence of exogenous L-DOPA. Thus, G. diazotrophicus expresses a
dioxygenase enzyme able to cleave the aromatic ring of L-DOPA in the same manner
as the 4,5-extradiol-DOPA-dioxygenases of plant origin. This allowed us to monitor its
elution by HPLC with diode array detection (DAD) and to determine the spectral
maximum wavelength of the peak at a λ of 470 nm. A minor peak with the maximum
λ of 405 nm and an Rt of 14.46 min was also detected, compatible with the presence
of betalamic acid. HPLC-ESI-TOF MS and a real standard of betalamic acid were used to
confirm its presence, with an exact mass of 212.0554 m/z.

G. diazotrophicus 4,5-DODA’s sequence, expression, and purification. A se-
quence with characteristics suitable to provide the dopaxanthin-forming activity de-
scribed to occur in cultures of G. diazotrophicus was found in its genome, the
WP_012222467 protein (gi 501179334), from which an optimized synthetic sequence
was expressed in Escherichia coli Rosetta 2 (DE3) cells. The DODA from G. diazotrophicus
(GdDODA) was expressed, and it accounted for 46% of the total soluble protein in the
cell extract. Recombinant protein was purified by Ni²⁺-chelating affinity chromatog-
raphy, subjected to SDS-PAGE, and purified to homogeneity (Fig. 1E), with a purification
Table of Contents

Bacteria Produce Betalains through a Dioxygenase Enzyme

Molecular and structural characterization of G. diazotrophicus 4,5-DODA. HPLC-ESI-TOF MS mass spectra showed a single peak with a molecular mass of 17.822 kDa, consistent with the molecular weight calculated using the protein sequence (17.8 kDa). Peptide mass fingerprint analysis further support the identity of this protein (Table 2; FIG 1).

**TABLE 1** Expression and purification of G. diazotrophicus dioxygenase

| Step                  | Vol (ml) | Protein (mg/ml) | Total protein (mg) | Activity (µM · min⁻¹) | Sp act (µmol · min⁻¹ · mg⁻¹) | Purification fold | Yield (%) |
|-----------------------|----------|-----------------|--------------------|-----------------------|-------------------------------|-------------------|-----------|
| Crude extractb        | 6.0      | 15.9            | 95.4               | 1.578                 | 0.595                         | 1.0               | 100       |
| Ni²⁺ chromatography   | 7.0      | 4.5             | 31.8               | 0.823                 | 1.087                         | 1.8               | 61        |

aActivity was determined using a 50-µl protein solution under the assay conditions.
bCrude extract was obtained from a cellular paste harvested from a 0.5-liter culture.
In addition, *G. diazotrophicus* 4,5-DODA was determined to be a dimer under native conditions after gel filtration because different samples from 2.5 μg up to 0.39 mg eluted as a single peak with a molecular mass estimated at 36.5 kDa (Fig. 1F). These results further support the sequence homology found for GdDODA. We determined a 46% identity (56.6% similarity, local alignment) (Fig. 1G) to a structurally characterized enzyme from *Burkholderia xenovorans* Lb400 (16, 17), a dioxygenase (Protein Data Bank [PDB] accession number 2NYH) that was also demonstrated to be a dimer, which is the closest homolog structurally characterized. By structurally assisted sequence comparison, GdDODA residues were assigned to specific secondary motifs (18, 19). A three-dimensional modeling of the novel enzyme from *G. diazotrophicus* was performed by the comparative modeling engine ProMod3 (20, 21) (see Text S1 [materials and methods] in the supplemental material) and then used for comparison with the only crystallized protein known to form betalamic acid from L-DOPA in enzyme assays, the protein YgiD (PDB accession number 2PW6) from *Escherichia coli*, a homologue of the plant enzymes (22) (Fig. 1H). Structural comparison (23) shows how both enzymes have common local structural features. Despite the protein size differences, the small monomer of GdDODA superimposes well with one portion of the YgiD protein, and the structures composed of the amino acids Gly28-Asp41, Val45-Pro57, His58-Thr59, Leu60-Ala66, Phe67-His83, and Gln84-Pro97 in the *G. diazotrophicus* sequence are also present in the plant homologue protein.

**Kinetic characterization.** The addition of the novel *G. diazotrophicus* enzyme to a reaction media containing L-DOPA produced a yellow coloration with a $\lambda_{\text{max}}$ of 414 nm (Fig. 2A). The optimum pH for the DOPA-dioxygenase activity was determined to be pH 6.5 (Fig. 2B), and this pH was used to characterize the kinetic parameters (Fig. 2C).

### TABLE 2 Main peptides identified to fully characterize the protein

| Peptide identified                      | m/z  |
|-----------------------------------------|------|
| (–)MTPVPEIRQITIGSYHAVHYFDPDG(DGR)      | 3,210.58 |
| (R)QIGTIGSYHAVHYFDPDG(DGR)             | 2,190.04 |
| (R)DIGWLGQPRALLGSR(L)                  | 1,638.91 |
| (R)DIGWLGQPR(G)                        | 1,041.55 |
| (–)MTPVPEIR(Q)                         | 1,039.56 |
| (R)AAIADR(F)                           | 616.34  |
| (R)DHLR(D)                             | 540.29  |

*Its peptide mass fingerprint (PMF) was determined by MALDI-TOF analysis after trypsin digestion. Amino acids in parentheses correspond to the theoretical residue after trypsin digestion. (–), the beginning of the sequence.*

---

**FIG 2** *G. diazotrophicus* dioxygenase activity characterization. (A) Spectral evolution of the transformation of DOPA (2.5 mM) by the addition of pure 4,5-DODA enzyme to the reaction medium. Spectra were recorded at 10-min intervals for 180 min, using a scanning speed of 2,000 nm–min$^{-1}$. (B) Effect of pH on dioxygenase activity. Reactions were performed with 2.5 mM L-DOPA in 50 mM sodium acetate buffer for pH values ranging from 3.5 to 5.5 and in 50 mM sodium phosphate for pH values ranging from 5.5 to 8.5. (C) Enzyme activity dependence on L-DOPA concentration measured in 50 mM sodium phosphate buffer, pH 6.5. AU, arbitrary units.
determined as a $K_m$ of 1.36 ± 0.31 mM and a $V_{max}$ of 5.26 ± 0.43 μM·min$^{-1}$. This $K_m$ value is lower than those obtained for Beta vulgaris 4.5-DODA (6.9 mM) (24), E. coli YgiD (7.9 mM) (22), and Amanita muscaria dioxygenase (3.9 mM) (25), making GdDODA the enzyme with the highest affinity for L-DOPA in the formation of the structural unit of betalains and the fastest one. The value for the turnover number was calculated as a $k_{cat}$ of 0.50 ± 0.019 min$^{-1}$, and the value for the specificity constant was calculated as a $k_{cat}/K_m$ of 0.36 ± 0.01 min$^{-1}·$mM$^{-1}$. Oxygen exchange was determinant in the production of betalains (Fig. S3). The production of dopaxanthin under inert atmosphere was negligible, while saturation with air provoked a production of dopaxanthin 50 times higher than that with the formation of the pigments when the exchange of oxygen was limited by diffusion. Substrates structurally related to L-DOPA, such as catechol, 4-methyl-catechol, and dihydrocaffeic acid, were also tested. These substrates showed kinetic parameters of $K_m$ 3 order of magnitude higher than the $K_m$ for L-DOPA (Table S1). In addition, theses alternative substrates showed a strong inhibition of the DOPA-dioxygenase of G. diazotrophicus by excess of the substrate, also called substrate inhibition (26). Inhibition curves and the kinetic mechanism are shown in Fig. S4. This makes L-DOPA the substrate most likely to be of physiological relevance. This is also supported by the presence of dopaxanthin, the final product of L-DOPA transformation in the culture medium.

**Mass spectrometry analysis of reaction products and intermediates.** The nature of the products derived from the enzymatic activity was analyzed by HPLC. The presence of betalamic acid and dopaxanthin described above for G. diazotrophicus cultures was confirmed (Fig. 3, peaks 3 and 5) (3). Additionally, a peak with an $R_t$ of 16.32 min and a $\lambda_{max}$ of 403 nm (Fig. 3, peak 4) was found in the reaction media, consistent with the product of 2,3-DOPA-extradiol-dioxygenase activity, muscaflavin. Two additional earlier peaks with a $\lambda_{max}$ of 361 nm were obtained with $R_ts$ of 8.04 and 8.54 min (Fig. 3, peaks 1 and 2) and identified as the 4,5-seco-DOPA and 2,3-seco-DOPA precursors. All products were characterized by ESI MS and TOF MS, confirming the nature proposed above (Table 3). Fragmentation spectra for all the reported compounds with annotations are provided in Fig. S5. The accurate ESI mass spectra showed the detection of molecular protonated ions, [M+H]$^+$, with exact mass values of 230.0665 m/z, 212.0562 m/z, and 391.1141 m/z (experimental masses). These results are consistent with the calculated masses for the molecules 4,5- and 2,3-seco-DOPAs (230.0659 m/z), with a difference of 2.52 ppm, betalamic acid and muscaflavin (212.0553 m/z), with a difference of 2.07 ppm, and dopaxanthin (391.1136 m/z), with a difference of 1.25 ppm. All the different values in this analysis are below the accepted accuracy threshold for elemental composition analysis, established at 5 ppm (27). For first time, these masses are experimentally determined for readily obtained seco-DOPA intermediates, and hence a complete and unambiguous picture of intermediates and final products of the evolution of L-DOPA in the presence of DOPA-dioxygenase is obtained.

**Chemical formation and comprehensive enzymatic-chemical mechanism.** The presence of the intermediates 4,5 and 2,3-seco-DOPAs reaches a maximum 5 h after the reaction started, and then these intermediates diminish until they disappear (Fig. 4A). Betalamic acid reaches its maximum value at 20 h and experiences a further decrease, while muscaflavin accumulates during the time in which the reaction was monitored. The formation of both betalamic acid and muscaflavin is preceded by a lag period, justified by the need for the formation of the corresponding seco-intermediates, before these molecules can be obtained. The lag period was 33 min for betalamic acid, and after its maximum, the concentration decrease was due to its condensation with intact L-DOPA molecules, resulting in the formation of dopaxanthin (Fig. 5), after a lag period of 3.35 h (Fig. 4A). Absolute concentrations of dopaxanthin, betalamic acid, and muscaflavin in the medium are provided in Fig. S6. The analysis of this chemical reaction yields a kinetic constant of 189 h$^{-1}·$M$^{-1}$. The timescale obtained (Fig. 4C) is in accordance with the results shown in Fig. 4A and justifies the time needed to obtain
color in developing flowers even when the initial enzyme-catalyzed reaction is finished. This chemical reaction implies a not previously considered additional time in the formation of the final pigments in betalain biosynthesis. It is a spontaneous chemical, but not immediate, reaction which depends on pH (Fig. 4C).

**TABLE 3** HPLC-ESI-TOF MS analysis of the reaction products formed by *G. diazotrophicus* dioxygenase activity in water supplemented with l-DOPA at 7.6 mM

| Compound        | Chemical formula | [M + H]+ (m/z) | Main-daughter ion (m/z) | Secondary-daughter ion(s) (m/z) | TOF exact mass (m/z) (exptl) | Calculated mass (m/z) (theoretical) | Δppm |
|-----------------|------------------|----------------|-------------------------|-------------------------------|-------------------------------|-----------------------------------|------|
| 4,5-Seco-DOPA   | C₆H₁₁NO₆         | 230.2          | 140.0                   | 187.1, 94.1                   | 230.0659                      | 230.0553                         | 2.07 |
| 2,3-Seco-DOPA   | C₆H₁₁NO₆         | 230.2          | 140.0                   | 94.1                          | 230.0659                      | 230.0553                         | 2.07 |
| Betalamic acid  | C₆H₉NO₅          | 212.0          | 166.1                   | 138.0                         | 212.0562                      | 212.0553                         | 2.07 |
| Muscaflavin     | C₆H₉NO₅          | 212.0          | 166.0                   | 149.0                         | 212.0562                      | 212.0553                         | 2.07 |
| Dopaxanthin     | C₁₂H₁₈N₂O₈        | 391.3          | 347.1                   | 301.1, 255.1                  | 391.1141                      | 391.1136                         | 1.25 |
The identification of the final product and intermediates makes GdDODA the smallest dioxygenase able to produce betalamic acid described in the literature (22, 24, 28–30). All betalamic acid-forming dioxygenases described up to now can be analyzed in terms of sequence homology and phylogeny (Fig. 1I). There is a compact clade formed by plant sequences (2) and the YgiD protein separated from A. muscaria dioxygenase. The dioxygenase from G. diazotrophicus stands apart from these two branches. An extended phylogenetic analysis was performed, and the results shown in Fig. S7 reveal the homology of G. diazotrophicus DODA with the sequences from Bradyrhizobium, Komagataeibacter, Mesorhizobium, and Inquilinus. All these sequences stand apart from those of previously characterized enzymes. Thus, sequence analysis reveals that GdDODA corresponds to a hitherto-unknown group of enzymes involved in this biosynthetic pathway.

Taking into account the evolution described and the nature of the intermediates characterized, the scheme shown in Fig. 4B is proposed for the “spontaneous cyclization” of the first compound in the biosynthesis of betalains. It shows how a nucleophilic addition with a concomitant proton transfer is possible and favored by the presence of the carboxylic group generated in the enzymatic cleavage of the ring. Water elimination by acid catalysis gives the structural unit of betalains. The same scheme can be...
In summary, an enzymatic-chemical mechanism is described for dioxygenase formation of betalamic acid and betalains which puts together the enzyme-catalyzed reactions and the evolution of the intermediates. The one-pot experiments described show how a single enzyme can produce the final betalain product in hours. The intermediates have been followed and structurally characterized thanks to an extraordinarily high activity detected in the enzyme from *G. diazotrophicus*, the first bacterium described able to synthesize betalains. The reactions, intermediates, products, constants, and lag periods described constitute the clearest experimental evidence of the first reactions involved in the route of betalain biosynthesis, expanded now to prokaryotes.

**MATERIALS AND METHODS**

**Chemicals, bacterial strains, plasmids, and enzymes.** Distilled water was purified using a Milli-Q system (Millipore, Bedford, MA, USA). HPLC-grade acetonitrile was obtained from Fisher Scientific UK (Leicestershire, United Kingdom). *E. coli* Rosetta 2 (DE3) cells, *E. coli* DH5α cells, and the plasmid pET28a were obtained from Novagen (Merck KGaA, Darmstadt, Germany). A HyperLadder 1-kb DNA ladder was obtained from Bioline Reagents Ltd. (London, United Kingdom). Restriction enzymes and the protein ladder were obtained from New England BioLabs (Ipswich, MA, USA). T4 DNA ligase was from Roche Diagnostics (Basel, Switzerland). *Pfu* DNA polymerase was purchased from Thermo Fisher Scientific Inc. (Waltham, MA, USA). A QIAprep spin plasmid miniprep kit, QIAquick PCR purification kit, and QIAquick gel extraction kit were from Qiagen (Hilden, Germany). All other chemicals and reagents were obtained from Sigma (St. Louis, MO, USA).

**Gluconacetobacter diazotrophicus culture.** *Gluconacetobacter diazotrophicus* was acquired from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ; the German collection of microorganisms and cell cultures) under accession number DSM5601. Active growing microorganisms were inoculated in 20-ml cultures of a specific medium that contains, per liter, 25 g D-mannitol, 5 g yeast extract, and 3 g peptone. These cultures were maintained overnight at 25°C with agitation. Afterwards, half of the cultures were centrifuged for 10 min at 5,000 × g and resuspended in 20 ml distilled water. The other half were kept in the culture medium. Both water and medium cultures were supplemented with various concentrations of l-DOPA (from 0 to 7.6 mM) and sodium ascorbate (15 mM) and further cultured at 25°C. Appropriate control media without l-DOPA were also cultivated at 25°C for each condition. After 2 days, the media were collected and analyzed by HPLC-ESI-TOF MS in a search for compounds derived from the synthesis of betalamic acid.

**Gluconacetobacter diazotrophicus DODA sequence and cloning.** The sequence of the protein WP_012222467, a hypothetic aromatic ring-cleaving dioxygenase from *Gluconacetobacter diazotrophicus* PAS5, has been deposited at the National Center for Biotechnology Information (NCBI, Bethesda, MD, USA) under code gi 501179334. This sequence was used as a template to synthetically obtain the 4,5-DODA.
sequence from G. diazotrophicus, enhanced for E. coli expression (GeneArt, Regensburg, Germany). PCR amplification was performed using Phusion DNA polymerase and the following primers, which include the restriction sequences recognized by the enzymes NdeI and XhoI: GdDODA-F (5’TATATACATATGACA CCGGTGCCGCGAA) and GdDODA-R (5’ATATATCTCGAGTTAACATGGGTTGC). This amplification yielded a 457-bp product, and its length coincides with that of the entire DODA synthetic gene plus the recognized sequences for the restriction enzymes. The PCR product was digested with NdeI and XhoI, purified with a QIAquick PCR purification kit, and inserted into the expression vector PET28a downstream of the T7 RNA polymerase promoter, the expression vector was previously digested with the same restriction enzymes. This produced the recombinant plasmid PET28a-GdDODA, which encodes an additional 22-amino-acid N-terminal sequence containing a 6 x His tag. The plasmid was transformed into E. coli DH5α (Novagen) electrocompetent cells and plated onto LB agar plates containing kanamycin. The resulting colonies were then analyzed by PCR. The plasmid PET28a-GdDODA was obtained from positive colonies using the Qiagen spin plasmid miniprep kit. The sequence was confirmed by DNA sequencing of the plasmid and subsequently used in further experiments.

Expression and purification. The GdDODA protein derived from plasmid PET28a-GdDODA was expressed in E. coli Rosetta 2 (DE3) (Novagen) and grown at 37°C in LB medium containing chloramphenicol and kanamycin to an A_{600} of 0.8 to 1.0. Induction was performed for 20 h with different concentrations of the inductor isopropyl-1-thio-β-D-galactopyranoside (IPTG) at different temperatures. Cells were harvested by centrifugation and resuspended in sodium phosphate buffer (50 mM, pH 8.0) with 0.3 M sodium chloride. Cell lysis was performed by sonication in a Cole-Parmer 4710 series ultrasonic homogenizer (Chicago, IL, USA). For trial scale and parameter optimization, chemical lysis was performed using BugBuster protein extraction reagent (Novagen). Recombinant protein was purified by His-select affinity gel (Sigma) according to the manufacturer’s instructions and then desalted using PD10 columns (GE Healthcare, Milwaukee, WI, USA) and eluted into Tris-HCl (20 mM, pH 8.5) buffer. Protein was quantified using the Bradford protein assay (Bio-Rad, Hercules, CA, USA) (31), and bovine serum albumin was used as the standard to obtain a calibration curve. Samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) by application to 12% polyacrylamide gels and stained using a standard Coomassie blue method.

Gel filtration. Samples of pure recombinant protein were applied to a Superdex 200 10/300 GL column equilibrated with sodium phosphate buffer (50 mM, pH 7.5) with 150 mM NaCl. The protein was eluted with the same buffer at a flow rate of 0.5 ml min^-1. Elutions were performed in an Äkta purifier apparatus (General Electric Healthcare) and monitored at 280 nm. Column calibration was performed with the following protein markers (Sigma): cytochrome c (12.4 kDa), carbonic anhydrase (29 kDa), alcohol dehydrogenase (150 kDa), and β-amylase (200 kDa).

MALDI-TOF MS protein analysis. The matrix solution for peptide analyses was α-cyano-4-hydroxycinnamic acid (20 mg ml^-1) in acetonitrile (CAN)-water-trifluoroacetic acid (TFA) (70:30:0.1). The peptide sample was dissolved in 0.1% TFA and mixed with the matrix solution. One microliter of this mixture was applied to the atmospheric-pressure matrix-assisted laser desorption ionization (AP-MALDI) target plate and allowed to dry. Experiments were carried out with an Agilent TOF mass spectrometer (Agilent Technologies, Santa Clara, CA, USA), equipped with an AP-MALDI ion source with an N2 laser (337 nm). Samples were measured in reflectron mode to identify molecular formulas based on precise mass measurements in positive mode. External calibration of the spectrometer was performed with standard peptides from the ProteoMass Peptide MALDI-MS calibration kit (Sigma). Data were recorded and processed with Agilent MassHunter Workstation software. Peptide mass fingerprinting was performed using Agilent Spectrum Mill software. Determination of protein absolute molecular mass and processed with Agilent MassHunter Workstation software. Peptide mass fingerprint determination standard peptides from the ProteoMass Peptide MALDI-MS calibration kit (Sigma-Aldrich, St. Louis, MO, USA). Two different peptides were used as controls (cytochrome c and carbonic anhydrase; Sigma-Aldrich). All data were recorded and processed through the Agilent MassHunter Workstation Qualitative Analysis Software (Agilent Technologies), and the intact molecular weight of the protein was obtained using the deconvolution algorithm from this software.

Trypsin digestion. The protein sample was prepared in 100 μl of buffer NH₄HCO₃ (50 mM, pH 8.0) with 0.02% ProteaseMAX surfactant (Promega, Madison, WI, USA). After that, the sample was reduced with dithiothreitol (DTT) at 10 mM and 56°C for 20 min and alkylated with iodoacetamide 50 mM at room temperature in the dark for 20 min. One microgram of proteomics-grade trypsin (Promega) was added, and the sample was incubated at 37°C for 4 h. Finally, the sample was centrifuged at 15,000 x g for 1 min to collect the condensate, and 0.5% TFA was added to stop the digestion. Peptides were cleaned up with C₁₈ ZipTips (Millipore) and evaporated using an Eppendorf vacuum concentrator, model 5301.

HPLC analysis of metabolites. A Shimadzu (Kyoto, Japan) LC-20AD apparatus equipped with an SPD-M20A photodiode array detector was used for analytical HPLC separations performed with a
A linear gradient was performed using water with 0.05% TFA as solvent A and acetonitrile with 0.05% TFA as solvent B. The linear gradient was performed for 24 min from 0% solvent B to 35% solvent B, the flux was 1 ml min⁻¹, and the column operation temperature was 30°C. Fresh samples were analyzed after enzymatic reactions. In the case of the time course for the enzyme reaction, 1.7 ml of the reaction medium was placed in an HPLC vial with the thermostatic block set at 25°C, and injections were directly performed at different times. Formation of pigments by the purified enzyme was also assayed under an inert atmosphere (nitrogen). Oxygen exchange effect was measured with and without aeration.

In this case, all samples contained 700 µl phosphate buffer (0.2 M, pH 6.5), 300 µl sodium ascorbate (100 mM), 1 ml l-DOPA (7.6 mM), 500 µl purified enzyme, and 500 µl MilliQ water up to a final volume of 3 ml. After 48 h, the samples were collected and analyzed by HPLC. In all cases, the injection volume was 50 µl.

Electrospray ionization mass analysis of metabolites. An Agilent VL 1100 apparatus with an LC mass selective detector (MSD) Trap was used for HPLC-ESI MS analyses. Elution conditions were analogous to those described above, using the same column. The vaporizer temperature was 350°C, and the voltage was 3.5 kV. Nitrogen at a pressure of 45 lb/in² was used as the sheath gas. Samples were ionized in positive mode. The ion monitoring mode was full scan in the range m/z 50 to 600. The electron multiplier voltage for detection was 1,350 V. A TOF–quantitative TOF (Q-TOF) Agilent 6220 MS equipped with a dual ESI-atmospheric pressure chemical ionization (APCI) interface was used for accurate mass determinations. Samples were ionized in positive mode, using a capillary voltage of 3.5 kV. Nitrogen was used as the drying gas, the gas temperature was 350°C, flux was set at 11 liter min⁻¹, and the nebulizer pressure was 40 lb/in². All data were processed through the MassHunter software (Agilent Technologies).

Absorbance spectroscopy. Enzyme activity was determined using a continuous spectrophotometric method by measuring the absorbance due to betalamic acid and muscalfavin appearance at a λ of 414 nm (24, 25). Unless otherwise stated, the reaction medium contained 50 mM sodium phosphate buffer, pH 6.5, 2.5 mM L-DOPA, and 10 mM sodium ascorbate (AA). Measurements were performed at 25°C in 96-well plates in a Synergy HT plate reader (Bio-Tek Instruments, Winooski, VT, USA). The final sample volume was 300 µl. The plate reader detector signal was calibrated with betalamic acid solutions of known concentration. The molar extinction coefficient at 424 nm, with an ε of 24,000 M⁻¹ cm⁻¹, was taken for the quantification of betalamic acid (33). Additionally, a JASCO V-630 spectrophotometer (JASCO Corporation, Tokyo, Japan) was used to measure the accumulation of reaction products over time, and the samples were carried out with a wavelength scan from 250 to 700 nm every 10 min for 3 h at 25°C. Measurements were performed in triplicate, and mean values and standard deviations were plotted. Errors associated with the results provided correspond to the residual standard deviations. Kinetic data analysis was carried out by using nonlinear regression fitting with SigmaPlot Scientific Graphing for Windows, version 10.0 (Systat Software, San Jose, CA, USA).

Chemical transformation of betalamic acid. The condensation of betalamic acid was kinetically characterized by determining the kinetic constant of the reaction. In the kinetic scheme for a second-order reaction in which two reactants yielded a single product following a stoichiometry of 1:1, the mass balance of the system has to be considered according to the following schemes: 

\[ k_\text{app}[\text{bet}] = \text{reaction rate} \]

\[ [\text{bet}]_0 \] is the initial concentration of betalamic acid before the reaction is triggered, and [dopax] is the concentration of dopaxanthin formed by the chemical reaction. By keeping constant the concentration of l-DOPA, a pseudo-first-order kinetics can be obtained as follows: 

\[ k_\text{app} = k[D\text{OPA}] \]

where \( k_\text{app} \) is the apparent constant at a fixed DOPA concentration. The substrate concentration equation 

\[ ([\text{bet}]) = ([\text{bet}])_0 e^{-kt} \],

where \( t \) is time, can be transformed into the product accumulation 

\[ ([\text{dopax}] = ([\text{dopax}])_0 (1 - e^{-kt})). \]

By representing the accumulation of dopaxanthin, [dopax], versus the initial concentration of betalamic acid, [bet]₀, the apparent constant can be determined from the slope of the linear representation obtained in the graph as \( k_\text{app} = (\ln \text{slope} - \ln 1)/t \), and then the kinetic constant of the second-order reaction can be determined.

**SUPPLEMENTAL MATERIAL**

Supplemental material for this article may be found at https://doi.org/10.1128/mBio.00345-19.

**TEXT S1**, PDF file, 0.1 MB.

**FIG S1**, PDF file, 0.1 MB.

**FIG S2**, PDF file, 0.1 MB.

**FIG S3**, PDF file, 0.1 MB.

**FIG S4**, PDF file, 0.1 MB.

**FIG S5**, PDF file, 0.4 MB.

**FIG S6**, PDF file, 0.1 MB.

**FIG S7**, PDF file, 0.5 MB.

**TABLE S1**, PDF file, 0.1 MB.
ACKNOWLEDGMENTS

We are grateful to José Rodriguez and to Alejandro Torrecillas (SAI, University of Murcia, Murcia, Spain) for skillful technical assistance in ESI-MS/MS, MALDI-TOF, and TOF/Q-TOF mass spectrometry experiments.

This work was supported by the Spanish Ministry of Economy and Competitiveness (MEC-FEDER) (grants AGL2014-57431 and AGL2017-86526) and by the Programa de Ayudas a Grupos de Excelencia de la Región de Murcia, Fundación Séneca, Agencia de Ciencia y Tecnología de la Región de Murcia (project 19893/GERM/15). L.E.C.-L. enjoyed a fellowship from Consejo Nacional de Ciencia y Tecnología (CONACYT, Mexico), and M.A.G.-R. holds a contract financed by MEC-FEDER (Spain).

We declare no competing financial interest.

REFERENCES

1. Gandía-Herrero F, García-Carmona F, Escrivano J. 2005. Floral fluorescence effect. Nature 437:334–336. https://doi.org/10.1038/437334a.

2. Gandía-Herrero F, García-Carmona F, 2013. Biosynthesis of betalains: yellow and violet plant pigments. Trends Plant Sci 18:334–343. https://doi.org/10.1016/j.tplants.2013.01.003.

3. Escrivano J, Pedreño MA, García-Carmona F, Muñoz R. 1998. Characterization of the antiradical activity of betalains from Beta vulgaris L. roots. Phytochem Anal 9:124–127. https://doi.org/10.1002/(SICI)1099-1565(199805)9:3<124::AID-PCA401>3.0.CO;2-4.

4. Gandía-Herrero F, Escrivano J, García-Carmona F. 2012. Purification and antiradical properties of the structural unit of betalains. J Nat Prod 75:1030–1036. https://doi.org/10.1021/jn200950n.

5. Sreekanth D, Arunasree MK, Roy KR, Chandramohan Rehdy T, Reddy GV, Reddanna P. 2007. Betanin a betacyanin pigment purified from fruits of Opuntia ficus-indica induces apoptosis in human chronic myeloid leukemia cell line-K562. Phytomedicine 14:739–746. https://doi.org/10.1016/j.phymed.2007.03.017.

6. Khan MI, Sri Harsha PSC, Giridhar P, Ravishankar GA. 2012. Pigment chemistry cell line-K562. Phytomedicine 14:739–746. https://doi.org/10.1016/j.phymed.2007.03.017.

7. Lechner JF, Wang L-S, Rocha CM, Larue B, Henry C, McIntyre CM, Riedl KM, Schwartz SJ, Stoner GD. 2010. Drinking water with red beetroot food color antagonizes esophageal carcinogenesis in N-nitrosomethylurea-treated rats. J Med Food 13:733–739. https://doi.org/10.1089/jmf.2008.0280.

8. Reyes-Izquierdo T, Pietrzkowski Z, Argumedo R, Shu C, Nemzer B, Wybraniec S. 2014. Betain-rich red beet concentrate improves reduced knee discomfort and joint function: a double blind, placebo-controlled pilot clinical study. Nutr Diet Suppl 6:9.

9. Guerrero-Rubio MA, Hernández-García S, García-Carmona F, Gandía-Herrero F. 2019. Extension of life-span using a RNAi model and in vivo cancer cell cytotoxicity of Rivina humilis L. berries, potential source of betalains. LWT Food Sci Technol 47:315–323. https://doi.org/10.1016/j.lwt.2012.01.025.

10. Lechner JF, Wang L-S, Rocha CM, Larue B, Henry C, McIntyre CM, Riedl KM, Schwartz SJ, Stoner GD. 2010. Drinking water with red beetroot food color antagonizes esophageal carcinogenesis in N-nitrosomethylurea-treated rats. J Med Food 13:733–739. https://doi.org/10.1089/jmf.2008.0280.

11. Segel IH. 1975. Enzyme kinetics: behavior and analysis of rapid catalytic processes. New York, NY: John Wiley & Sons, Inc. 400 p.

12. Smith TF, Waterman MS. 1981. Identification of common molecular subsequences. J Mol Biol 147:195–197. https://doi.org/10.1016/0022-5396(81)90087-5.

13. Von-Ardenne R, Döpp H, Musso H, Steglich W. 1974. Isolation of muscaflavin by recombinant dopa-dioxygenase from Amanita muscaria. Phytochemistry 30:169–174. https://doi.org/10.1016/S0031-9422(91)84119-D.

14. Gloor P, Aebi U, Zbinden C, Zbinden M. 1991. Biogenesis of betalains: purification and partial characterization of xanthine dehydrogenase from Glottiphyllum longum. Phytochemistry 30:169–174. https://doi.org/10.1016/S0031-9422(91)84119-D.

15. Segel IH. 1975. Enzyme kinetics: behavior and analysis of rapid equilibrium and steady state enzyme systems. Wiley-Interscience, New York, NY.

16. Smith TF, Waterman MS. 1981. Identification of common molecular subsequences. J Mol Biol 147:195–197. https://doi.org/10.1016/0022-5396(81)90087-5.

17. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Siquizzato S, Park YM, Buseo N, Lopez R. 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res 43:W580–W584. https://doi.org/10.1093/nar/gkv279.

18. Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaei B, Kedouas V, Notredame C. 2006. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Res 34:W604–W608. https://doi.org/10.1093/nar/gkl092.

19. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENGrid server. Nucleic Acids Res 42 W320–W324. https://doi.org/10.1093/nar/gku316.

20. Biasini M, Bienter S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefeler F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252–W258. https://doi.org/10.1093/nar/gku340.

21. Bienter S, Waterhouse A, De Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T. 2017. The SWISS-MODEL Repository—new features and functionality. Nucleic Acids Res 45:D313–D319. https://doi.org/10.1093/nar/gkw1132.

22. Gandía-Herrero F, García-Carmona F. 2014. Escherichia coli protein YqiD produces the structural unit of plant pigment betalains: characterization of a prokaryotic enzyme with DOPA-extradiol-dioxygenase activity. Appl Microbiol Biotechnol 98:1165–1174. https://doi.org/10.1007/s00253-013-4961-3.

23. Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaei B, Kedouas V, Notredame C. 2006. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Res 34:W604–W608. https://doi.org/10.1093/nar/gkl092.

24. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENGrid server. Nucleic Acids Res 42 W320–W324. https://doi.org/10.1093/nar/gku316.

25. Biasini M, Bienter S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefeler F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252–W258. https://doi.org/10.1093/nar/gku340.

26. Bienter S, Waterhouse A, De Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T. 2017. The SWISS-MODEL Repository—new features and functionality. Nucleic Acids Res 45:D313–D319. https://doi.org/10.1093/nar/gkw1132.

27. Ferrer I, Garcia-Reyes JF, Fernandez-Alba A. 2005. Identification and quantitation of pesticides in vegetables by liquid chromatography time-of-flight mass spectrometry. Trends Anal Chem 24:671–682. https://doi.org/10.1016/j.trac.2005.04.004.

28. Mueller LA, Hinz U, Zryd J-P. 1997. The formation of betalamic acid and muscaflavin by recombinant dopa-dioxygenase from Amanita. Phytochemistry 44:567–569. https://doi.org/10.1016/0031-9422(96)00625-5.

29. Christinet L, Burdet F, Zaiko M, Hinz U, Zryd J-P. 2004. Characterization and functional identification of a novel plant 4,5-extradiol dioxygenase.
involved in betalain pigment biosynthesis in Portulaca grandiflora. Plant Physiol 134:265–274. doi.org/10.1104/pp.103.031914.

Sasaki N, Abe Y, Goda Y, Adachi T, Kasahara K, Ozeki Y. 2009. Detection of DOPA 4,5-dioxygenase (DOD) activity using recombinant protein prepared from Escherichia coli cells harboring cDNA encoding DOD from Mirabilis jalapa. Plant Cell Physiol 50:1012–1016. doi.org/10.1093/pcp/pcp053.

31. Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. doi.org/10.1016/0003-2697(76)90527-3.

32. Gandía-Herrero F, García-Carmona F, Escribano J. 2005. A novel method using high-performance liquid chromatography with fluorescence detection for the determination of betaxanthins. J Chromatogr A 1078:83–89. doi.org/10.1016/j.chroma.2005.05.013.

33. Trezzini GF, Zrÿb J-P. 1991. Characterization of some natural and semisynthetic betaxanthins. Phytochemistry 30:1901–1903. doi.org/10.1016/0031-9422(91)85036-Y.

34. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242–W245. doi.org/10.1093/nar/gkw290.